



How to Collect and Analyze Data

A Manual for Sheriffs and Jail Administrators

—Second Edition—

National Institute of Corrections

Morris L. Thigpen, Director

Virginia Hutchinson, Chief
Jails Division

Jim T. Barbee, Project Manager

How to Collect and Analyze Data

A Manual for Sheriffs and Jail Administrators

—Second Edition—

by

Gail Elias
Voorhis Associates, Inc.

September 1999

© 1997 by Voorhis Associates, Inc.

These materials were originally produced under National Institute of Corrections grant #EE-8 awarded to Gail Elias, Principal, Voorhis Associates, Inc., and were revised through funding from the National Institute of Corrections, U.S. Department of Justice. The National Institute of Corrections reserves the right to reproduce, publish, translate, or otherwise use and to authorize others to publish and use all or any part of the copyrighted material contained in this publication.

Notice: This document contains references to specific computer software programs and published materials. Inclusion of product names in this document does not imply endorsement by the National Institute of Corrections or the U.S. Department of Justice, nor does exclusion of others imply non-endorsement.

Table of Contents

PREFACE	vii
Who Should Use This Document—and Who Will Not Need To	vii
What This Manual Will—and Will Not—Do	vii
A Word About Good Detective Work	viii
ACKNOWLEDGMENTS	ix
INTRODUCTION	x
The Situation	x
Traditional Resources Have Failed to Help	x
Resource 1: Introductory Statistics Courses	xi
Resource 2: Information Systems Seminars	xi
Resource 3: Research Methods	xii
Additional Resources	xii
CHAPTER 1: GOOD MANAGEMENT REQUIRES GOOD INFORMATION	1
What Is Management?	1
Ways to Get the Facts	2
What Can You Do with Data?	3
Example 1: Better Budgeting and Allocation of Funds	3
Example 2: More Effective Deployment of Staff	4
Conclusion	5
CHAPTER 2: INFORMATION THAT SHOULD BE COLLECTED	6
Writing Good Problem Statements	6
Identifying Data Elements from Problem Statements	7
A Catalog of Correctional Data Elements	8
Inmate Population Data Elements	9
Inmate Profile Data Elements	10
Operational Data Elements	14
Criminal Justice System Performance Data Elements	16
A Word About Collecting Data in Other Agencies	19
Conclusion	19
CHAPTER 3: PREPARING FOR THE DATA COLLECTION	20
Data Collection Skills	20
Where to Find the Skills You Need	22
Working with People from Outside the System	23
Conclusion	24

CHAPTER 4: HOW TO LOCATE AND CAPTURE INFORMATION	26
Jail Information Sources	26
The Problem with Data Sources—and Some Solutions	32
Problem 1: Data Elements Are Scattered	32
Problem 2: Data Is Missing	32
Problem 3: Information on the Form Is Wrong	33
Problem 4: Data Elements Are Poorly Defined	34
Problem 5: The Handwriting is Illegible	34
Problem 6: No One Keeps Data for the Whole System	35
Setting Up an Information System for the Jail	35
Elements of the System	35
System Start-Up	35
Special Issue Data Collections	44
Designing Special Data Collection Forms and Code Books	45
Conclusion	51
 CHAPTER 5: HOW TO PUT IT ALL TOGETHER	52
Doing the Data Collection	52
Determine an Overall Strategy	52
Determine How Much Data to Collect	52
Coding	56
Rules for Coding Data	57
Coding in an Automated System	58
Gather Your Data Collection Supplies	61
Begin	61
Automation Issues	62
Conclusion	63
 CHAPTER 6: HOW TO ANALYZE INFORMATION	64
What Is Statistics?	64
Statistically Significant Differences	65
Probability	65
Normal (Bell-Shaped) Distribution	66
Statistical Results	67
Types of Statistics	68
Descriptive Statistics	68
Statistics for Examining Differences Between Groups	75
Statistics to Examine the Relationship Between Data Elements	78
Statistical Sins	81
Sin 1: How Could the Results Be Wrong If the Calculations Were Perfect?	81
Sin 2: What Is the Real Average?	82
Sin 3: What Is Missing from this Picture?	82
Sin 4: So What If It Is Statistically Significant?	83
Sin 5: If A, Then B?	83
Sin 6: What Is Wrong with this Picture?	84
Conclusion	85

CHAPTER 7: HOW TO INTERPRET INFORMATION	86
Case Study One: The Friday Night Blues	86
The Situation	86
Before Going on to Case Study Two	99
Case Study Two: The Alternative Answer	100
The Situation	100
Conclusion	114
CHAPTER 8: SHARING INFORMATION WITH OTHERS	115
Why Pictures Are Worth a Thousand Numbers	115
Methods for Displaying Data	116
Tables	116
Barcharts	117
Pie Charts	118
Line Graphs	119
Guidelines for Good Graphics	119
Conclusion	126
CHAPTER 9: GETTING THE MOST FROM YOUR INFORMATION SYSTEM	128
What This Chapter Is—and Is Not—About	128
A Short Course in Computers	128
Overview of Computer Elements	129
The Central Computer	129
Peripherals	130
Communication Devices	130
Overview of Software	131
Applications Software	131
A Last Word on Hardware and Software	133
Getting the Most Out of a Data Base	133
What to Do Before Buying or Upgrading a Computer	134
Determine the Purpose of the Information System	134
Define Your Needs	135
System Expectations and Guidelines	139
The User Manages the System	139
The System Avoids Duplication of Effort	139
The System Is Reliable	139
The System Has a Safety Net	139
Match the Right Application to the Task	140
Know How the System Handles Updated Information	140
Format Flexibility	140
Determine Internal Checks	140
User Guidelines	141
Explore the System	141
Practice Analysis	141
APPENDIX A: A GLOSSARY OF STATISTICAL TERMS FOR NON-STATISTICIANS	142

APPENDIX B: ANNOTATED BIBLIOGRAPHY	149
Resources for the Criminal Justice System	149
Other Resources	152
APPENDIX C: MANUAL DATA COLLECTION PROCEDURES AND SAMPLE FORMS	155
Elements of the System	155
System Start-Up	155
APPENDIX D: INMATE PROFILE DATA COLLECTION	165
Data Collection Form	165
Inmate Profile Data Collection Code Book	167
APPENDIX E: INCIDENT DATA CODE BOOK	174
APPENDIX F: TRANSPORT DATA COLLECTION	177
APPENDIX G: TABLES FOR DETERMINING SAMPLE SIZE	180
APPENDIX H: SIMPLE RANDOM SAMPLING	183
Types. Methods. and Procedures of Sampling	183
APPENDIX I: CALCULATING THE STANDARD DEVIATION	191
APPENDIX J: CALCULATING CHI SQUARE	194
APPENDIX K: MANUAL DATA DISPLAY	199
Barcharts	199
Pie Charts	200
Line Graphs	201
Graphic Tools	202
APPENDIX L: AUTOMATED CASE STUDY GUIDE	204

Preface

Who Should Use This Document—and Who Will Not Need To

Although the title of this document, *How to Collect and Analyze Data: A Manual for Sheriffs and Jail Administrators*, indicates it was written for individuals who are in policymaking roles in corrections, it does not mean others in the chain of command—from lieutenants to line officers—will not also find it helpful. Its concepts and content also apply to community corrections facilities, juvenile detention facilities, and other institutions. It will also be beneficial to people outside the jail who are involved in the collection and analysis of jail data. Actually, the manual should be helpful to anyone who has to gather information about jail problems, policies, and practices. *How to Collect and Analyze Data* addresses both ongoing and special issue data collections in local jails.

Those who have been involved in in-house data collections or who have crime analysts or statisticians on staff will find that many of the items covered in this manual are familiar. This manual is **not** intended for criminal justice policy analysts or planners who regularly work with statistics, information systems, or techniques like systems analysis. However, because of its focus on the issues facing jail administrators, it may be a useful resource for policy analysts or statisticians who are new to the criminal justice arena.

It is particularly difficult to address issues associated with automation in this document. Computer hardware and software are changing at an incredible rate. Whatever is said here today will probably be outdated tomorrow. While most jails and local law enforcement agencies have computers, **what** they have and **how** they use them **vary** tremendously. Some restrict computers to National Crime Information Center (NCIC) or office functions, like word processing. Many have "jail systems," but their capabilities are so varied that it is impossible to generalize. As a result, unlike the first edition of this manual in which it was easy to assume that there were few automated jails, this manual assumes that you have some degree of automation, but makes no further assumptions about what. Rather, the manual focuses on what can be done with off-the-shelf software. For many small jails, some combination of this software may be all that is needed to get into the "information system" business. For larger jails, some combination of this software may be the easiest way to begin the process of turning the jail information system into a management information system.

What This Manual Will—and Will Not—Do

This manual cannot be all things to all people. It will not teach you everything you ever needed to know about statistics, but it will make you a more informed **consumer** of the statistics you receive. It will also provide a basic explanation of common descriptive statistics written in understandable language, not statistical jargon. It will not teach you how to design the most up-to-date computerized management information system, but it will help you identify the elements that should be included in such a system

and will provide some help in getting started. It will not make you a criminal justice policy analyst, but it will give you the opportunity to analyze some real data and show how that information was used in policy decisionmaking in other criminal justice systems.

This manual has several goals:

- To identify reasons why sheriffs and jail administrators should collect data;
- To "de-mystify" statistics and data collection procedures;
- To provide tips on how to collect data in the simplest, easiest, most efficient way possible that allows policymakers to draw valid conclusions;
- To provide an opportunity to practice analytical skills; and
- To provide guidelines for displaying the information collected clearly and effectively so that county officials, other agencies, and the public can understand it.

A Word About Good Detective Work

I hope that you find this manual to be a useful tool, which introduces you to a new way of understanding and solving problems. Statistics are tools that managers can use to "get the facts." That is all analysis really is. Analysis is very much like good detective work. It involves looking at a series of facts and trying to piece together a picture of what has happened.

Those with a background in law enforcement actually have a head start on many individuals who are just beginning to try to analyze information. Many of the skills are the same. You will probably have a theory about the things you want to analyze. You will need to look for facts that support or refute that hypothesis. You will still need to "play your hunches." The tools that you will be using may be strange at first, but once you are used to them, they will be as comfortable as the techniques you use to solve cases. I hope this manual makes you feel more at home with statistics and data and that it encourages you to apply these tools to solve the problems that your jail and criminal justice system face.

Gail Elias

Acknowledgments

In a sense, many people contributed to this manual, both academics (who tried to teach me something about statistics, policy analysis, and research methods) and practitioners (who taught me more than I ever wanted to know about the "nitty-gritty" of collecting data in jails). To all of you—who know who you are—thanks for your help and your patience.

Introduction

The Situation

At times, correctional organizations appear to be slowly sinking under the weight of their own paperwork. Like law enforcement and the courts, jails document tremendous amounts of information.

- Law enforcement officers complete paperwork associated with arrests and complaints.
- Court personnel document countless appearances and dispositions in addition to the text of every hearing.
- Jail staff record details, such as descriptions of inmate property, operational procedures, head counts, incidents that occur while an individual is in custody, etc.

Considering the vast amounts of time involved in these activities, it is frustrating and disheartening to think that so little of the information captured can be easily retrieved later for use in management decisionmaking. However, this state of affairs is not too surprising when you think about the purposes of both manual and automated recordkeeping systems found in criminal justice agencies. These information systems are designed to store data about people or events; information is put together on a case-by-case or event-by-event basis. It is relatively easy to retrieve information about a specific person or event. However, it is often difficult to **aggregate** that information to answer a question for management.

The way in which data is identified for each information system often makes it impossible to look for specific data in any other way than by searching through the entire recordkeeping system, file by file. In the case of automated systems, often information is lost when staff "write over" one of the variables. For example, when an inmate's bond is reduced, the original bond amount may be lost. Many older information systems used data structures that did not link data together in a way that made it easy (or in some cases possible) to combine data from multiple files or sources. These technical problems have frustrated many data collection efforts.

Traditional Resources Have Failed to Help

When sheriffs and jail administrators tried to learn how to collect jail data or set up a management information system in the jail, many found that the traditional resources available to them were not very helpful. These resources included:

- Statistics course work,
- Information systems seminars, and
- Research methods.

Resource 1: Introductory Statistics Courses

Most statistics courses focus on mathematical foundations of statistics, **derivations**, and computation, which are things most correctional managers do not need to know. Computers can perform these computations for you. And memorizing statistical formulas does not make a whole lot of sense, since they are readily available in texts. Yet frequently, that is what success in statistics course work requires.

In addition, most statistics courses require breaking through a statistical language barrier. Not only do you have to deal with mathematical formulas, but you also have to deal with phrases like "95% confidence intervals," "significant at the .001 level," and "non-parametric tests of significance," which mean little at first exposure to statistics. And the explanations of these **terms** are often worse than the terms themselves!

In the long run, introductory statistics classes do not deal well with **analysis**, which is precisely what sheriffs and jail administrators hope to get from these courses. Analysis is the process of figuring out what something means and what, if anything, can be done about it. Statistics are tools that can aid in that process, not the process itself. Sheriffs and jail administrators want to apply these tools to "real-life" problems. Instead, in most statistics classes, they get a series of instructions on how to make the tool. As a result, beginning statistics courses leave most sheriffs and jail administrators frustrated, confused, or yawning. This manual avoids statistical jargon, focuses on statistical assumptions, and exposes sheriffs and jail administrators to data analysis and its associated statistics.

From time to time, this manual will explain the statistical term for the concept being discussed. This may be necessary so you are prepared when a serious student of statistics starts talking about "standard deviations" and "sampling schemes." Any statistical terms introduced are defined in the text and then summarized in Appendix A.

Resource 2: Information Systems Seminars

Information systems seminars often focus on computer hardware and software and forget about data. In the excitement of the computer "bells and whistles" and the complex terminology, it is too easy to assume that a particular type of equipment will solve all of management's information problems. Many sheriffs and jail administrators attend these seminars to determine what jail management software they should buy. That is, unfortunately, often putting the cart before the horse. To evaluate how good an automated system is and how well it will meet your needs, you first need to know what data you should capture.

Actually, the process for designing a manual system is a precursor for designing an automated system. However, many software companies and designers do not emphasize this point, because they have an investment in selling their computer hardware and software. This may turn out to be a sizable additional cost when programs have to be tailored to meet the needs of a specific department. Sheriffs and jail administrators who do not have the resources to invest in a specialized jail information system may hope they can buy "off-the-shelf" software and a personal computer and program it to meet their needs. However, problems in learning to use the applications, setting up the procedures to ensure that staff enter data consistently, and training people how to use the software frequently make these individuals reconsider. And then, for many small organizations, there is the sizable issue of making the commitment to actually enter the data that is needed.

Frequently, too, the language associated with information systems is even more confusing than that associated with statistics. Conversations are spiced with words like "data base," "LANs," and others. None of this seems to help much when all you really wanted to know is how to set things up so you could find out how many pre-trial inmates charged with felonies were held for more than a month in 1990 by some easier method than looking at all the prisoner files and the shift activity sheets for that year. In frustration, many sheriffs and jail administrators throw up their hands and continue gathering information like they had in the past, or they do not collect the information at all.

This manual provides assistance in designing a manageable manual information system and transferring key management information system elements to a personal computer using off-the-shelf software. There are many ways to develop this system. Two strategies discussed in this manual are:

1. Extracting data from an automated information system to a personal computer and manipulating the data using off-the-shelf software, and
2. Designing a manual information system to collect data that may not be available in an operational system and manipulating the data using off-the-shelf software.

Resource 3: Research Methods

Sheriffs and jail administrators who try traditional research or program evaluation course work may actually have more of their specific questions answered than those who try the previously mentioned resources. This type of course work will provide some answers about how to define problems, design questionnaires, and organize data collections.

However, several problems attach to this resource as well. Research methods also have a language of their own, and the complexity of research designs that require control and experimental groups may seem unrealistic to sheriffs and jail administrators. The jail's often competing missions of safety, security, and service do not mesh well with traditional research models. As a result, there is a tendency to "throw out the baby with the bath water" and disregard useful aspects of these courses because other aspects make very little sense to sheriffs and jail administrators faced with real decisions and real problems. This manual will provide "real world research methods that can help with jail management.

Additional Resources

Since this manual was first published in 1982, other materials that deal with data collection have been published. Many of these publications are both practical and pertinent to correctional settings. Although a number are written in technical jargon and others intimidate by their very size, there are more available resources now. Appendix B contains a bibliography of data collection manuals and materials with some comments regarding their contents. It also lists several statistics and evaluation texts, as well as other books on the subject. These should be helpful to those getting started collecting and analyzing data and those who are organizing an information system.

Despite the fact that a considerable amount of material has been developed over the last decade, there is still a surprising lack of materials and resources for those who need basic information about how to collect data in their correctional facilities. This manual will help fill that gap by identifying:

Why you should collect data,

- How to identify the information that is needed,
- How to organize a data collection,
- How to locate and capture that information,
- How to put the data collection together,
- How to analyze the information,
- How to interpret the results, and
- How to share that information most effectively

This manual also includes the following resources you may find useful in collecting data:

- A list of common data elements that jails should begin to collect,
- A description of the skills needed to collect data,
- A list of common places where data elements are found in jails,
- A model manual information system,
- A model information system constructed using off-the-shelf software,
- A sample data collection sheet, code book, and a corresponding computer format in which to process this information,
- A short introduction to descriptive statistics,
- A series of programmed learning exercises to practice your analytical skills, and
- Examples of charts and graphs that can be used to display data.

These resources can be adapted to meet the needs of individual jails.

Chapter 1: Good Management Requires Good Information

In the past, managing jails was considered to be so basic that there was nothing to it. However, with increased court intervention in correctional matters, demands for better management of correctional facilities increased in the 1970's. The courts discovered that often the difference between a "constitutional jail" and an "unconstitutional jail" was the way in which the facility was managed. Since good management relies strongly on good information, sheriffs and jail administrators found that their organizational world had become a much more complex place in which to work.

As a result, "professional management" arrived in correctional facilities. Sheriffs and jail administrators were introduced to a number of techniques, such as cost benefit analysis, Total Quality Management (TQM), and organizational development (OD), that were designed to help them improve organizational performance. Sometimes these techniques were very helpful, and sometimes they were not. In analyzing their relative successes and failures, the ability of the organization to generate good, valid information about its problems emerged as a critical variable.

What Is Management?

Management is mostly about mobilizing an organization's resources, in this case the jail's, to solve or avoid problems. In many cases, problems are not solved in the sense that they go away, but the organization finds a way to manage them more effectively. Regardless of the complexity of the problem, a basic management formula applies to analyzing a situation. It consists of five steps:

1. Facts are gathered.
2. Facts are interpreted in light of the organization's mission and values.
3. Alternative solutions are developed.
4. A decision is made.
5. Action is taken.

All of the steps in the formula are critical to its successful application. Facts are the foundation on which everything that follows depends. If the facts are wrong, you are likely to reach the wrong conclusion, make the wrong decision, and take the wrong action. Good detectives do not bring a case to court unless they have gathered the evidence. Correctional managers are in precisely the same position. They should avoid making policy decisions without the best available information. This manual provides guidelines for getting the facts.

Ways to Get the Facts

Facts can be gathered in two major ways:

- By quantitative methods and
- By qualitative methods.

Quantitative methods rely on numbers to reveal something about an event or phenomenon. They usually focus on questions like, "How often?" or "How much?" Qualitative methods describe an event or phenomenon. Law enforcement professionals use qualitative methods all the time. Two of the most commonly used qualitative methods are:

- Interviews and Observation.

Qualitative methods are excellent tools for problem solving. However, even in the hands of trained observers and interviewers, some problems attach to using them. First, as human beings, we have certain perceptual biases; we hold certain values and assumptions about the way the world operates. And we process all information in the light of those biases and assumptions. Qualitative tools provide "checks" on these biases.

Another problem associated with qualitative measures lies in the fact that "where we sit often determines where we stand." All officials in the criminal justice system have specific sets of tasks that they perceive as their "mission." Together with the expectations that other individuals have of them, these tasks comprise their roles. Because of their roles, different criminal justice officials often see the same event differently.

For example, someone charged with major narcotic sales may post a bond. The court has no problem with this as long as the defendant comes to court. A crowded jail may see the individual as a relatively good risk for release since he/she is not violent. However, law enforcement sees things differently because they suspect that this individual will probably disappear. If a researcher approached the judge, the jail administrator, and the chief of police regarding this case, with questions about changing bonding procedures, they would have decidedly different reactions. With qualitative methods, the researcher would get a great deal of rich, descriptive information, but there would be no way to evaluate how often such a situation actually occurs or how often the individual actually absconds.

Without quantitative data, there is no way of knowing how many people bond out of jail, the level at which bond is set, how many fail to appear, and many other questions that should be answered when bonding criteria are reviewed. Information gathered in numerical form is more objective and limits individuals' perceptual biases in interpreting the information. This is not to say that statistics cannot be biased, because they can be. However, it is rarely the numbers that are biased. Most often, it is their interpretation or representation that is slanted. Chapter 6 addresses these issues and how to guard against them.

Perhaps the greatest advantage of gathering information in quantitative form is that you can determine how much error is involved. Not only can you estimate what the chances are of being wrong,

but you can also estimate **how** wrong you are likely to be. In other words, you can identify how many times out of a hundred you would come up with a similar answer and how much the new answer would probably vary from the first one.

In addition to focusing on the **facts** portion of the management equation, this manual focuses specifically on quantitative ways of getting facts. This is the process we call collecting, analyzing, interpreting, and displaying data.

What Can You Do with Data?

Sometimes the hardest thing about collecting data is convincing yourself that it can be useful. Two examples of how jail data can be used to make better management decisions follow.

Example 1: Better Budgeting and Allocation of Funds

For many jails, budgeting is a somewhat mystical process, in which the next year's decision is based on what was spent last year. This figure is then divided into four equal amounts for each of the four quarters. Sometimes this process "backfires" when additional funding without any unusual circumstances is required because critical changes were not considered. Let's consider one area of the budget—food service—and apply the traditional approach as well as one that is somewhat different.

Last year, two small, nearly identical facilities with average daily populations of 50 prisoners each spent \$63,875 for food. Operating on the "let's ask for 5% over last year's" approach, Sheriff #1 requests \$67,069 for food in the next year's budget. Sheriff #2 decides to try a different approach.

The local Consumer Affairs Office advises Sheriff #2 that inflation hit food products heavily last year, causing an increase of 5% in food costs, and that costs are projected to increase by at least 5% next year. Sheriff #2 has also been following increases in the jail population closely. The statistics on "Average Daily Population," "Length of Stay," and "Jail Days" have also increased regularly. Over the past 5 years, Jail Days have increased by about 5% each year. Furthermore, Sheriff #2 knows that the jail was fuller from July through December.

Last year, Sheriff #2 calculated that \$3.50 was spent on food each day for each prisoner in the facility. At 5% inflation, it is anticipated that \$3.67 will be necessary to provide the same meals next year. And, if the population continues to increase, Jail Days will increase to approximately 19,175 for the whole year. As a result, rounding the totals off, Sheriff #2 asks for \$70,375 for food.

Let's assume that Sheriff #2 is also a good politician and that the county commissioners provide the full amount requested. Instead of dividing the money into four equal amounts of \$17,593 and change per quarter as the accountant recommends, Sheriff #2 calculates how much money will really be used each quarter if the jail population behaves as it has over the last 5 years. Usually, 19% of the Jail Days fall during each of the first two quarters of the year; 29%, during the third quarter; and 33%, during the fourth quarter. So, by dividing the \$70,375 by those percentages, Sheriff #2 **knows** that about \$13,371.25 should

'Jail Days is the total number of days spent by all prisoners in jail during a fixed period of time, such as a month or year.

be spent during each of the first two quarters of the year, \$20,408.75 during the third quarter, and \$23,223.75 during the last quarter.

Based on this scenario, assuming Sheriff #2 can stick to the budget, the funds should last through the fiscal year without any difficulty. Sheriff #1 will have spent all the funds in the food budget somewhere around the first week of December.

Similar strategies can be used to calculate future costs of any items supplied to the inmate population. A little information and a relatively simple mathematical procedure helped Sheriff #2 make a more accurate estimate of future expenditures and monitor the rate at which dollars were being allocated for a specific function. It also got Sheriff #2 a slightly larger budget.

Example 2: More Effective Deployment of Staff

Jails have traditionally been understaffed. If personnel are among the jail's most precious resources, it is critical that they are deployed in the most effective manner. Let's assume that Sheriff #1 and Sheriff #2 are evaluating their staffing patterns. Both know that their average daily populations are about 50 and that there are about 3,500 bookings a year (about 9.5 bookings per day).

Sheriff #1 looks at these statistics and decides to staff the Booking Room 365 days a year, 24 hours a day. After all, prisoners are brought in 7 days a week, 24 hours a day, and personnel have to staff the Booking Room. If the shift relief factor is 1.75, 5.25 officers will be needed to cover each post in the Booking Room. Sheriff #1 decides that two officers are needed. This results in the need for 10.5 officers. When presented with this information, the Board of Commissioners expresses great concern about the costs attached to this staffing pattern and refuses to give Sheriff #1 the money to hire additional staff.

Sheriff #2 decides that more information is needed before deciding how to deploy staff. Because Sheriff #2 suspects that more people are booked during the weekend, a search for data to validate this assumption begins. Sheriff #2 has the jail administrator go over the Booking Log for the previous year and prepare a chart that shows the number of bookings on each shift for the last year by day of the week. All the jail administrator has to do is count the number of bookings that took place on each shift each day of the week. The results are displayed in Table 1.1.

Table 1.1 Yearly Bookings by Shift and Day of Week (1995)

Shift	Mon	Tues	Wed	Thurs	Fri	Sat	Sun	Total
0:00-8:00	115	75	75	90	45	350	350	1,100
8:00-16:00	95	85	55	115	85	225	150	810
16:00-24:00	115	125	100	175	450	425	200	1,590
Total	325	285	230	380	580	1,000	700	3,500
Average/day	6.25	5.48	4.42	7.30	11.15	19.23	13.46	9.50

Sheriff #2 decides against full-time booking officers on all but the 16:00-24:00 shift when the volume of activity is heavy enough to warrant a full-time position. During the week on the other shifts, jail staff assigned to other areas should be able to handle incoming prisoners. However, Friday nights and weekends are very busy in the Booking Room. Sheriff #2 elects to post an additional staff member in the booking area on Friday evenings and weekends with instructions that the "roving jail officer" assigned to each shift will assist in the Booking Room when it is busy. Ultimately, this results in the equivalent of

2.29 officers posted in the Booking Room—one 7 days a week on the swing shift, and a second officer on all shifts from 16:00 on Friday through 8:00 on Monday. When the shift relief factor is considered, Sheriff #2 needs a total of four officers for the Booking Room. And the sheriff believes that the area will be covered more effectively during peak periods of use than it had been in the past.

Conclusion

The examples given above are only two of many that illustrate how good information can be used to help sheriffs and jail administrators manage more effectively. Generating and using good data is a critical aspect of dealing with such jail issues as:

- Crowding,
- Facility planning,
- Policy and procedure development,
- Controlling violence in facilities,
- Managing special inmates,
- Transporting prisoners,
- Cost-effective operations,
- Staffing, and
- Determining appropriate programs and services.

These examples suggest that a lot of good can come from collecting, analyzing, and using good information. In addition, most standards for correctional facilities mandate that jails maintain records systems containing information that can (and should) be used for management purposes. In some cases, management information systems are even specified. Furthermore, compliance with standards requires a considerable amount of documentation. As discussed in Chapter 4, much of the documentation you keep to defend against litigation will be part of the management information and data collection system.

Each sheriff or jail administrator may decide **why** to collect data and **how** to use it in his/her own situation. Information is power. It is up to you to tap it, channel it, and put it to good use.

Chapter 2: Information That Should Be Collected

Many people find it difficult to identify all the information (or data elements) needed to improve both the quantity and quality of information available for management decisionmaking. As a result, sheriffs and jail administrators may rely on professional analysts to define what will be collected. The problem is that the analysts will select data elements that they think jails could use, and most computer software designers or system analysts know relatively little about jail operations.

This chapter will help you decide which data elements should be collected on a routine basis. It also provides guidelines for developing good problem statements, since they are the basis for identifying the data you will need to gather in special **issue** data collections. It suggests how organizational or program goals can be used to identify data elements. Finally, it categorizes and lists data elements commonly collected in jails and identifies other criminal justice data elements that should be of concern to the jail although they may not be readily available in the jail itself. These data elements will probably become part of your management information system.

Writing Good Problem Statements

This section describes a process for identifying the data elements to be gathered in special issue data collections. Statements that describe problems in detail are actually composed of many basic data elements that should be collected; others may be suggested by the problem statement. Writing a good problem statement sounds elementary. However, most people find it difficult to describe a situation in its most basic terms, which is what a good problem statement does. Many people substitute solutions, obstacles, reasons, symptoms, and goals for good problem statements.

As shown in Table 2.1, each substitution, or imposter, has an associated cost.

Table 2.1 The Costs of Poor Problem Statements

Imposters	Costs
Solutions (ways of solving problems).	Other options that may be longer lasting or more cost-effective may be glossed over, or you may fail to solve the "right" problem.
Obstacles (conditions that make it difficult to solve problems).	Problems may seem insurmountable and no effort to change the status quo is made.
Reasons (why things happen).	It is possible to derail the whole process into a disagreement over who is right.
Symptoms (visible signs that indicate a deeper problem is present).	It is too easy to miss the factors that are really causing the problem.
Goals (the end results to be achieved).	No description of the current situation is developed, making it difficult to identify the data elements.

A good problem statement is a concrete description of the situation. It makes it possible to identify the data elements needed to determine causes and suggest solutions. For example, the following paragraph is a good statement of the crowding problem in a specific jail.

For the last year, the number of people who have been arrested and booked at the jail has steadily increased; their length of stay has increased, too. As a result, I have had prisoners sleeping on mattresses on the floor. Something has changed the way the criminal justice system works.

This is a useful problem statement, because it concretely describes the crowding that is troubling the jail. In the next section, this problem statement is used to identify the pieces of information that should be collected to determine what is causing the problem.

Identifying Data Elements from Problem Statements

Turning a problem statement into the data elements related to it is usually not too difficult. The better the problem statement, the easier it will be. Generally, by asking yourself the questions for which you are trying to find answers, the pieces of information you need begin to emerge. It may help to think of possible questions like:

- What has changed?
- How much?
- How often?
- Why?
- What is related to this change?

The problem statement written above directs the analyst toward pieces of information that will help unravel the reason for crowding. These include data elements like:

- The ones mentioned in the problem statement (i.e., number of people arrested and booked and length of stay),
- Charges on which persons are arrested (this year compared with last year),
- Numbers of summonses and citations issued by local law enforcement agencies (this year compared with last year),
- Amount of time from first appearance to arraignment, trial disposition (this year compared with last year),
- Average length of sentences (this year compared with last year), and
- Charges for which sentences are given (this year compared with last year).

While these data elements will not identify the ultimate cause of the problem, they will identify the arena in which the investigation should proceed. A good follow-up to this analysis would include interviews with key actors in other law enforcement agencies, the courts, and the prosecutor's office to identify any policy changes that would lead to crowding in the facility. This problem statement is good for two other reasons:

- It is based on knowledge of the criminal justice system and an understanding of the basic factors that lead to crowding (e.g., increased admissions and length of stay), and
- It includes an educated guess (a "hypothesis" in research terminology) about what may be causing the problem.

If the problem statement actually defines goals, the task of identifying the data elements becomes a bit more difficult, but it is still manageable. In fact, most program or performance evaluations start from a statement of the particular program's goals. The trick with goals, which are usually abstract statements, is to find something measurable that will be present if the goal is achieved. To put this in research terminology, you must look for measurable objectives that lead to the goal or an operational definition of the goal.

For example, suppose that Sheriff #2 has established a goal of providing a safe environment for staff, inmates, and visitors to the facility. After making a number of policy and procedure changes, Sheriff #2 decides to find out how well the organization is meeting that goal. Sheriff #2 and the jail administrator work together to define the conditions that describe a safe environment. These include the following:

- There are no assaults on staff, inmates, or visitors;
 - No contraband enters the facility;
 - No fires occur in the facility;
 - There are no accidental injuries to staff, inmates, or visitors; and
- The tension level in the facility is low.

Sheriff #2 and the jail administrator have no trouble identifying the data elements for most of these objectives. The data elements are:

- The number of assaults on staff, inmates, and visitors;
- The number of times that contraband was found either in the facility or before it entered the facility;
- The number of fires that occurred in the jail;
- The number of compensation claims;
- The number of accidental injuries recorded in the inmate medical records; and
- The number of disciplinary reports.

While the process of identifying data elements from goal statements is somewhat more complex than working from a problem statement, jails can identify the data elements needed to evaluate program performance. A number of good handbooks that can help sheriffs and jail administrators with this particular kind of data collection are listed in Appendix B.

So, to determine the data elements you need to collect, you will generally follow these steps:

1. Define the problem and its component parts.
2. Determine the data elements that are represented in the problem.
3. Identify potential causes and contributing factors.
4. Write the questions for which you are trying to find answers.
5. Determine the data elements represented in the potential causes, contributing factors, and questions.

A Catalog of Correctional Data Elements

The process described in the preceding section is helpful when collecting data that deals with a specific issue or problem. However, it makes sense to collect some data on a routine basis so that it is

readily available. This section categorizes, identifies, and describes frequent uses for commonly collected data elements. Some of these data elements should be routinely collected to be most useful; others will more likely be collected only when specific problems confront the jail.

Correctional data elements can be divided into four categories:

1. Inmate Population Data Elements,
2. Inmate Profile Data Elements,
3. Operational Data Elements, and
4. Criminal Justice System Performance Data Elements

Inmate Population Data Elements

Inmate Population Data Elements are called "population" data elements because information is kept about each person booked at the jail (the whole population in a statistical sense). In the language of information systems, they are summary numbers or statistics assembled by time period, typically a month or a year. While they can be calculated or recreated from most data bases, sometimes that may require more effort than you are willing or able to expend at the time you really need to know what the average daily population of your jail was in March last year (after that data has been archived) or what the difference between the average population and the peak population was two years ago. These data elements should be kept regularly; they are the items that should be included in a jail information system! Inmate population data elements include the following:

1. Average Daily Population (ADP) is an average of the number of people housed at the jail (in statistical terms, the mean). ADP is usually expressed as a monthly or yearly average. It is a computed data element and is calculated for one year by dividing total Jail Days by the 365 days in the year.* ADP is most often used as a way of expressing how full the facility is.
2. Length of Stay (LOS) is the number of days between a prisoner's admission and release. LOS is often expressed as an average. That way it refers to how long (on the average) each prisoner stays at the jail. LOS can also be expressed as a frequency distribution. In that case, the amount of time each prisoner stays is divided into groups (e.g., prisoners released within 24 hours, 48 hours, 72 hours, 1 week, etc.), and then the number of prisoners in each group is counted and expressed as a percentage (e.g., 66% are released within 24 hours, another 14% are released within 48 hours, etc.). LOS is almost always expressed in days for jails and in months for prisons. For pre-trial prisoners, it is used as a way of assessing the speed of the court process.
3. Jail Days is another measure of the amount of time spent in the jail. Jail Days are calculated by adding the time each prisoner spends at the facility within a given period of time (usually a month). In many respects, this data element is one of the most useful because ADP and LOS can both be calculated from it if the number of bookings and the amount of time it spans are known. Jail Days can be translated directly into the percentage of capacity at which the jail operates and the number of beds that are needed to accommodate a certain jail population, and it is often used with other data elements to develop cost ratios (e.g., food cost per inmate per day). Jail Days are

*An alternate way to calculate the average daily population is to record the number of inmates in the jail each day at the same time(s). Typically most jails should use an early morning count, when the institution is fullest. Average the counts to obtain the ADP.

often called other names, including prisoner man days, inmate days, resident days, days of care, etc.

4. **Total Bookings** is the total number of persons arrested and brought to the jail and is a good measure of Booking Room activity. Together with Total Arrests, Total Bookings suggests something about law enforcement practices in the jurisdiction (i.e., if and how often field citations are used). A word of caution is in order here, however. Sometimes persons are brought to the jail, detained there for some time, and not booked. These could include persons such as those held on mental or alcohol detainees. If these practices occur in your facility, you will need to keep separate statistics on these individuals.
5. **Net Bookings** is the number of persons actually housed at the jail after booking. Net Bookings also may be called inmates, prisoners, residents, etc. While Total Bookings is the best measure of activity in the Booking Room, Net Bookings is the best measure of activity in the jail itself. If you subtract Net Bookings from Total Bookings, the result will be the number of persons released directly from the Booking Room without ever entering general population. These individuals are released on recognizance, bond, a summons or citation, or at the authorization of the arresting officer.

Most jails maintain records on several subgroups of this data element as a way of describing the jail population. Common subsets of Net Bookings include:

- Felony and misdemeanor inmates, traffic offenders, and holds;
- Inmates who are pre-trial, post-trial, or on a hold or detainer;
- Male and female inmates; and
- Adult and juvenile inmates.

Jails frequently keep information on these subsets because of statutory requirements for separation. These statistics become particularly important when planning for a new facility because they are directly related to the number of beds that must be provided for certain types of prisoners.

These are the basic data elements that jails should routinely keep to describe the inmate population. They can be used for many purposes, including monitoring population levels in the facility, long-range facility planning, population forecasting to determine anticipated bed space needs, budgeting, deployment of staff, and program development.

Inmate Profile Data Elements

The following list represents many common inmate profile data elements. However, many additional data elements related to the unique needs of a specific jurisdiction deserve consideration. You will have to identify these for your jurisdiction on your own. You may elect to collect some of this information on a regular basis and other data elements only as some special issue or problem emerges. Your challenge is to determine which is which and how to quickly and easily retrieve those data elements that you do not routinely collect. Inmate profile data elements include

1. **Legal Status** describes the legal reason the prisoner is at the facility. General categories used with this data element are: pre-trial, post-trial, sentenced, and holds. This is one of the most basic descriptive data elements and is used in a variety of ways. Some uses relate to the kinds of programs in which the inmate may participate (e.g., only sentenced inmates may work); others

relate to the development of alternatives to incarceration or changes in arrest or court practices to relieve crowding. If your classification system is based on objective criteria, legal status is probably one of them. Be aware that legal status changes over time. An inmate may enter the jail in pre-trial status and leave in sentenced status. He/she may be pre-trial on a local charge, but on a hold for another jurisdiction. As a result, jurisdictions need to consider how to record this information. Jurisdictions may:

- Record legal status several times (e.g., at the time of admission and the time of release),
- Associate legal status with a specific charge, and/or
- Track legal status as it relates to the most serious charge.

Note: *Do not be surprised by what your automated jail management system does with this information. Some automated systems "lose" the initial data when information is updated or replaced. For example, if you had just one variable for Legal Status and you gathered information about people who left, you would not be able to tell how many people came in pre-trial and left sentenced because only the last variable which should have been updated when they were sentenced—will show up.*

2. Charge Status describes the level of charges on which the prisoner is detained or sentenced. General categories used with this data element are felon, misdemeanor, municipal ordinance violator, traffic offender, or hold (especially probation and parole violations). This is another very basic descriptive data element. Charge Status is frequently used as one of the criteria to decide if an individual is housed in maximum, medium, or minimum security.
3. Specific Charge or Charges is what the prisoner is charged with. This data element is often used to try to predict behavior within the facility. Charge at arrest is often a good indicator of who will remain in custody. Charge can be a particularly difficult data element because it changes over time. The charge at the time of arrest is often modified by the prosecutor. Plea negotiation may result in a third possible charge, for which the individual is sentenced. You may want to add a data element or elements that tell what the charges were at release and what the charges were on which the inmate was sentenced.

Inmates may be arrested on multiple charges. One strategy commonly used to deal with multiple charges is to select the "most serious offense" for the analysis. However, sometimes the "less serious" charges provide insight into the prisoner that may be very useful for other purposes, such as classification. For example, a prisoner charged with a felony residential burglary may also be charged with misdemeanor-level possession of a controlled substance. One strategy to deal with this issue is to develop one or two specialized charge data elements that consider all of the charges. These could include a data element that identifies whether any drug or alcohol offenses were included, a data element that identifies whether there were any FTA (failure to appear) or FTC (failure to comply) warrants or detainers, or any other category of charges that might be of specific local interest.

A word of caution in categorizing charges is in order. If you hope to pick out information about persons charged with specific offenses, all individuals charged with driving under the influence [DUI], you have to code your information at that level. You can group charges in a number of ways in later analyses (e.g., all traffic offenses or all alcohol-related offenses), but if you code the charges as a traffic or alcohol-related offense, you can never separate out the information for only DUIs. Chapter 5 contains more information about this coding procedure.

While charge information has to be entered at this level of detail so that data about specific offenders, such as DUI or domestic violence offenders, can be used to identify specific cases, it is often useful to categorize charges to look at groups, such as prisoners who are charged with crimes against persons, property offenders, public order offenders, etc.

Automated systems sometimes use a code, such as a statute or UCR (Uniform Crime Report) number, to record this information. Depending on the characteristics of the program, this can work well or create some interesting problems. The issue is whether the computer system checks the numbers your staff enter. If it does not, the potential for human error can produce some bizarre or non-existent charges on reports the computer generates. The computer system should have an internal validation process that alerts staff to the charge they enter or, better yet, have a system that is menu-driven that lets the operator pick the charge.

Sometimes automated systems let staff freely enter information for a variable. In data base terminology, these are "free-form" fields. The problem here is not visible until you try to analyze this data by computer.

Keep in mind that "Possession of a Controlled Substance," "Controlled Substance-Possession," and "Drug Possession" are the same charge. However, the computer will not understand that they are the same thing. Computers are very literal. This makes for a long and tedious process to recode information. If your computer program has free-form fields, a great deal of time will be saved if you ensure that staff always enter information the same way. Even better, use an alternative coding strategy, such as statute number, to avoid this problem.

4. Date of Birth is preferable to Age as a data element. Statistical studies have shown that people make more mistakes (or misrepresentations) when asked their age. Many computer programs can calculate Age from Date of Birth. This data element is often expressed as the average (mean) age of jail inmates. It is also useful to use another statistic, the median, to explore age. See Chapter 6 for more information about the mean and the median. Age is frequently more useful when displayed as a frequency distribution. This data element is often used to make decisions about the types of programs that would be appropriate for the inmate population (e.g., younger prisoners are more likely to use an active recreation program than older prisoners).
5. Ethnicity refers to the prisoners' racial or ethnic backgrounds. Common general categories include Caucasian, African-American, Asian, Native American, and Other. Hispanic and non-Hispanic are two other categories often included separately in this data element. The Ethnicity data element can be used in a number of ways. It may suggest cultural and ethnic groups from which detention personnel should be recruited, cultural or social customs or practices of which

staff should be aware, and potential language barriers. It may also be useful in dealing with the assumptions the community holds about who is in jail.

It is interesting that, categories make up 100%. For example, in a number of well as for census and crime respondents are characterized as Hispanic/Non-Hispanic separately from ethnic background.

6. **Education** is several potential data elements. They include the last grade completed, actual level of educational attainment as indicated by testing, interest in obtaining a general equivalency diploma (GED), interest in a remedial program, or interest in college course work. Its main application is in deciding what type of education program is appropriate for the inmate population.
7. **Vocational Skills** consists of several data elements, including skills attained and interest in skills building. This data element, in combination with LOS data, is useful in developing programs and services for the inmate population. It also provides information that is potentially useful for those designing a "bridge" or "reentry" program for inmates.
8. **Employment** is several data elements, including the individual's employment status at time of arrest (i.e., employed full-time, part-time, unemployed), last date employed, whether the prisoner is or was receiving unemployment benefits or other public assistance, and occupation (categorized). This data element is frequently used as one criterion for release on recognizance, and may also suggest programs or services that might be helpful to the prisoner in preparation for release (e.g., employment counseling).
9. **Family Ties** consists of several data elements, which include marital status, next of kin, residence, and number of dependents. All of these are potentially useful in assessing whether the prisoner is a good candidate for release on recognizance.
10. **Residence** is related to family ties and employment. It is frequently examined in considering release on recognizance.
11. **Physical Health Status** consists of several data elements, including current illnesses or injuries, disabilities, and history of physical illnesses. These data elements define the most likely medical conditions the medical staff will have to treat. They are also useful in preparing training for staff in first aid, cardiopulmonary resuscitation, etc., and in preparing and defending the medical budget, including requests for medical staff.
12. **Mental Health Status** consists of several data elements, including past or present treatment for mental health problems, type of treatment (e.g., inpatient vs. outpatient), whether a mental health crisis worker had to see the prisoner on intake, and whether special housing is required because of a psychiatric condition. These data elements are becoming more important to sheriffs and jail administrators as increasing numbers of seriously disturbed individuals wind up in jails. This information is useful in making decisions about the level of programs and services required by the inmate population and in identifying the types of personalities staff should be trained to manage. It is also helpful in pre-architectural programming, as these individuals have special

housing requirements. Other actors in the criminal justice system and the community are also interested in this information.

13. Substance Abuse Status consists of several discrete data elements, including the presence of past or present substance abuse problems, types of substances abused, whether there is a history of treatment for substance abuse, whether substance abuse is related to the reason the prisoner is in custody, whether the prisoner was under the influence when the crime was committed or under the influence when arrested and booked, and interest in treatment for substance abuse problems. These data elements are useful in making decisions about the kinds and levels of programs and services that should be provided for the inmate population and can also be used to identify areas in which staff will require training. They are also of interest to many other actors in the criminal justice system.
14. Release Status is a series of data elements that describe how the individual was released from the facility, including the location to which the prisoner was released; his/her legal status at the time of release (i.e., pre-trial vs. post-trial); if pre-trial, the bond type and amount (if a bond release) or release on citation; and, if post-trial, whether the prisoner was released for "time served," transferred to another facility, or charges were dismissed. These data elements can be particularly helpful in assessing how the local criminal justice system is working and in making choices about the appropriateness of alternatives to incarceration for some of the jail population.

Inmate Profile Data Elements contain some of the most useful information that can be gathered in the jail. They can be used to:

- Assess the impact (or potential impact) on the jail population of alternatives to incarceration and additional release options,
- Illustrate the impact of criminal justice policy and legislative decisions on the jail,
- Determine the causes of crowding,
Design and implement a classification system, and
Develop a pre-architectural program based on needs and describes the new facility's functions and programs.

Many jails do not collect all Inmate Profile Data Elements on a routine basis, although some of the data elements are almost certainly recorded in every jail. Data of this type is often collected at specific times to provide information about specific problems and is often collected for only a specially selected portion (a sample) of the inmates at the facility. There are a variety of ways in which samples can be selected in jails. Information about sampling is provided in Chapter 6.

Particularly in smaller jails, this is some of the most difficult data to gather because of the way most jail record systems are structured. It is also some of the most powerful information you can have. But unless your record system is atypical, you will not find a single source for the information you need, and much information can be lost when an inmate is discharged. If you are to capture this information, most likely you will need to fill out a special data collection sheet while the individual is in custody.

Operational Data Elements

Operations are central to any correctional facility. To monitor operational performance, Operational Data Elements are necessary to:

- Complete a staffing analysis,
- Change operational policies and practices,
- Evaluate organizational efficiency and effectiveness,
- Assess the "climate" in the facility, and
- Defend against litigation.

These data elements also can be grouped in "sets" that relate to specific operational areas. Some of the most common operational data elements follow.

1. **Transport Set** consists of data elements that provide information about transports to court and to other locations. The elements include the destinations of transport, reasons for transport, number of officers required for the transport, number of inmates transported, time in, time out, transport expenditures (e.g., meals, gasoline), and miles traveled. These data elements can be combined to identify the staff hours used (from which personnel costs can be calculated), the most frequent destinations, and the number and cost of medical transports (useful in assessing whether providing a medical service in-house would be more cost-efficient than transporting to hospitals, doctor or dentist offices, etc.). This data is also critical in supporting requests for additional staff.
2. **Activity Set** consists of data elements that provide information about routine activities, such as hours of recreation provided, the number of inmates who participated in recreational programming, the number and types of visits provided, the hours various types of programs are provided, the number of inmates served by various types of programs, etc. These data elements, like the transport data elements, can be linked to staff time to identify the cost of programs and services. Elements in this data set may be useful in grant applications, at budget time, and in documenting compliance with specific standards.
3. **Incident Set** consists of data elements that provide information about the crimes and incidents (significant events) that occur in the jail, including reason for incident; number of inmates involved; number of staff involved; location of incident; time of incident; date of incident; action(s), including disciplinary hearings, taken as a result of the incident; character of the incident (violent or nonviolent); and injuries and/or property damage resulting from the incident. You should already be documenting these data elements in a systematic way to use if legal action is taken against the jail. By analyzing these data elements, you may be able to identify problems that lead to incidents by identifying specific places, times, days, or staffing patterns when the incidents occur. Corrective action can be taken if patterns are identified.
4. **Personnel Set** summarizes all personnel changes and practices, including the number of sick days used per employee; vacation days; compensatory time authorized and taken; other leave time (e.g., military, maternity) authorized and used; rate of turnover, both by those leaving the department and by those vacating a position through transfer or promotion but remaining in the department; and demographic information (age, sex, race, etc.). This is another set of data elements that you already record in your personnel records. They are critical for establishing your shift relief factor (how many people it actually takes to fill one post that operates 24 hours a day, 365 days a year). They are also critical for addressing employee turnover, which is an ongoing problem in many jails.
5. **Cost or Budget Set** consists of data elements associated with the cost of running the jail, including all budget information, particularly categorical expenditures. Your budget or

accounting system should divide the dollars you are authorized into categories, such as personnel, fringe benefits, food, medical, inmate supplies, etc. Most budgets provide a considerable amount of detail about how the jail's money is spent.¹ These data elements are very useful when translated into ratios with Inmate Population Data. Examples of these ratios include the cost per inmate per day for meals, medical care, and staff, and the cost per inmate for photo processing (mug shots) and other functions. They are often used to evaluate the efficiency and effectiveness of jail operations.

Operational data elements are most frequently used for in-house purposes: to evaluate and monitor organizational performance, locate organizational trouble spots before they get out of hand, and document current problems that can then be rectified. However, these data elements have some obvious uses outside the organization, primarily in developing effective budget requests. In these days of limited and declining government resources, good budget justifications are critical. This type of data can help you prepare a well-documented rationale for funding.

Criminal Justice System Performance Data Elements

The policy choices of the other elements of the criminal justice system (primarily law enforcement and the courts) determine the size and nature of the jail population. Law enforcement practices determine the number of people who come into the criminal justice system in general and the jail in particular. On the other hand, the court (including the prosecutor, probation, the defense, and judges) determines the jail population by controlling the amount of time inmates spend in the facility. This leaves the jail with little ability to regulate the jail population.

The jail's influence over the size of the inmate population is indirect at best. As a result, it is essential to delve into the practices of the **other** elements of the criminal justice system to understand what is happening with the jail population. This is particularly important when planning a new facility and when trying to reduce crowding. Many data elements reveal a great deal about the policy and practices of other components of the system.

Law enforcement data elements. Data elements that are helpful in assessing the activity of law enforcement agencies follow.

1. Police Agency identifies the agency or police department responsible for the arrest. It is generally used to determine which agency accounts for what percentage of the bookings at the jail, and on what charges. Comparisons are often made from agency to agency or against a common set of arrest criteria.
2. Arrest Type attempts to document how the agency uses the jail. General categories of Arrest Type could include "book and releases" or station house bookings (in which prisoners are "booked and prints and photographs done prior to immediate release from the jail upon signing a citation), warrant arrests, and on-view arrests. This data element is often used in conjunction with Charges and other Inmate Profile Data Elements to determine if there are individuals being jailed who are candidates for release on citation or diversion alternatives.

¹For data analysis purposes, you need detail in the budget categories to track jail expenses separately from other costs. However, depending on the “rules” in your jurisdiction, it may be better to group expenses by categories. Make sure you can track internal costs in enough detail to make the information useful!

3. Summons and Citation in Lieu of Arrest refers to the number of summonses and citations written by law enforcement agencies. It is most commonly used in conjunction with Charges to determine if there are individuals who might be diverted from the jail.
4. Traffic-Related Arrests is used to identify arrests associated with a traffic violation in which a summons is not used. It is generally used in conjunction with Charges and Arrest Type to determine if there are individuals who are candidates for diversion from the jail.
5. Alcohol-Related Arrests identifies whether alcohol abuse was related to the arrest. It is generally used in conjunction with Charges, Arrest Type, and Inmate Profile Data Elements to determine if placement in a detoxification facility is a viable alternative to arrest.

These data elements can help determine which agency is the chief "supplier" of the jail's population and assess how each department uses the jail. They can also help assess the viability of diversion programs and the use of summons and citation as ways of reducing the jail population. This is especially true when the data identify the amount of time that individuals who are potential diversion candidates actually spend in the jail. Such information will help you assess whether the increased use of summons and citation (or other options) will significantly reduce crowding or the need for additional bed space.

Court data elements. Court data elements are critical to managing jail populations. They are often the hardest to collect because they are kept in another record system, which, like the jail's, is seldom organized to facilitate data collection. Since the court affects the amount of time that prisoners spend in the jail, Length of Stay and Jail Days become even more important when trying to assess the impact of changes in court policy and practice. Furthermore, it is critical to be able to identify the Length of Stay or the number of Jail Days of the group you are trying to divert. The average LOS for the whole jail will only give you misleading information. Thus, either LOS or Jail Days information must be collected with the court information. The court data elements follow.

1. Bondine Set is data relating to the bonding practices of the criminal justice system, including the types of bonds used in the jurisdiction, offenses for which bond is granted, the level at which bond is set, and the number of bond reductions. These data elements are often used with the Inmate Profile Data Elements, especially Employment Status, Residence, and Marital Status. Frequently, two other elements are added: Number of Prior Arrests and Number of Prior Failure to Appear. The primary application of these data elements is to determine if changes in bonding practices (e.g., the use of release on recognizance bonds or reduction in bond) would reduce the jail population,
2. Failure to Appear (FTA) Rate measures the number of people who were granted bond and later fail to return for court (also potentially the number of bond revocations) in comparison to the total number of bonds and citations given. Local criminal justice systems should determine what an acceptable FTA Rate is for their jurisdiction. This data element can be used to compare FTA Rates for individuals released through different mechanisms (e.g., those released with a street summons vs. those released after screening at the jail).
3. Offenses for Which Jail Sentences Are Given identifies the charges that resulted in sentences at the local jail. This data element is often used in conjunction with other data elements (e.g., Judge or Court) to assess differences in sentencing practices. If your jurisdiction has sentencing guidelines, this data element would be useful in assessing the degree to which they are followed.

4. Type of Release describes how inmates are released (e.g., cash bond, own recognizance, sentence served, etc.). Qualitative information about the system is needed to help interpret this data element
5. Court of Jurisdiction identifies the court with responsibility for the case. This data can be useful in comparing the processing efficiency of various courts. Different jurisdictions do not use the jail in the same way. If changes in practices are needed, it is critical to identify who and what must change. Both Court of Jurisdiction and Judge require some "special handling" to be useful. The special handling requires analyzing this data in the context of charge, prior offenses, FTA's and other violations, and other related variables.
6. Judge identifies which judge in a court has responsibility for the case. This data element should be handled with great care, because it can raise very sensitive issues around judicial discretion.
7. Defense Attorney and Prosecutor identify the attorneys involved in the case. Defense Attorney can be categorized as public defender, other appointed counsel, retained counsel, or none. Attorneys frequently manage cases differently and require different amounts of processing time. Changes in processing speed can be critical to reducing crowding at the jail. This data element is needed to identify points at which extra manpower or a change in policy could influence the jail population.
8. Offenses on Arrest, Filing, and Sentencing are all charges. "Charge" is a tricky data element to interpret even though its definition seems relatively clear. It needs to be interpreted in the context of information you have about criminal justice policy and practice. Some of the issues that emerge around this data element include:
 - Does law enforcement incorrectly charge the persons they arrest?
 - Does law enforcement "over-charge" the persons they arrest?
 - Does the prosecutor engage in plea bargaining?
 - Do law enforcement and the prosecutor charge bargain?

Often, the trickiest part of this data element is its political sensitivity. Use caution when you present information on this data element.

9. Sentence describes the sentence imposed by the court. It should include all alternative sentences used in the jurisdiction and the length of the sentence.
10. Critical Events are the dates of critical events in the court process and their outcomes. Since the length of time needed to process a case is the major way in which courts affect the jail population, it is critical to identify the amount of time each aspect of the process requires. The critical event data elements include the dates of arrest, first appearance, arraignment, trial, and sentencing. This data set should also include Number of Continuances (and on whose request). Outcomes of the major events should be categorized to identify what happened. Data elements should allow you to differentiate between cases in which a plea was entered, a trial actually occurred, and those dismissed (with and without prejudice).

Both law enforcement and court data serve a critical purpose in planning. To establish an appropriate level of confidence in any planning or forecasting done based on jail statistics, it is important to consider

how much the behavior of the other parts of the system has varied. Law enforcement and court data elements provide the best means of assessing the degree to which the behavior of other parts of the criminal justice system has fluctuated over the years and determining the actual variations in their practices.

A Word About Collecting Data in Other Agencies

Develop a strategy to make good use of any data you plan to collect in another criminal justice agency. First, the information you need to collect is not your information; the other agency owns it. The owner may view your interest in this information somewhat skeptically. Most sheriffs, for instance, would wonder why a chiefjudge was interested in average daily jail population data. Second, all record systems have their own quirks, and you are not likely to know what another agency's are. As a result, you can easily misinterpret the information. Third, significant technical issues are involved with getting different types of computers and computer programs to communicate with each other.

Many data collections that involve multiple agencies can benefit from a collaborative approach in which all of the agencies involved participate. In addition to addressing issues about the motivation for the data collection and problems of interpreting and accessing information, a collaborative approach tends to build commitment to both the results of the analysis and the solutions that will come out of the process. When key participants are **not** included in the data collection, the outcome is very predictable—they are likely to discount the data, the process by which it was gathered, or the analysis. In these situations, the participants tend to disagree about the data and not address the issue that is at the root of the problem.

Conclusion

Sometimes the hardest part of any activity is getting started. Knowing what can be collected and how other jurisdictions have used these data elements should help sheriffs and jail administrators decide what data elements they need to collect to solve specific problems. In many cases, the data elements discussed will more than cover the spectrum. Some systems will not require more than the Inmate Population Data Elements, a few of the Inmate Profile Data Elements, and the Operational Data Elements. Other systems that are confronting specific problems will need to do a special data collection that includes both Inmate Profile Data elements and the Criminal Justice System Performance elements.

You may sometimes be confronted with specific issues or problems that were not addressed here. You should define the data elements needed, using the methods suggested in the earlier section, Identifying Data Elements from Problem Statements. The program evaluation and research texts cited in Appendix B can provide some additional help. By now, you may be wondering when you will find the time to do a data collection and, if you hire someone, how much it will cost. For those who are not concerned about those issues, go directly to Chapter 4. For those who are concerned about those issues, see Chapter 3 for information on data collection strategies.

Chapter 3: Preparing for the Data Collection

When the data elements to be collected have been identified, the next big task is to decide how to collect them. This is not the same thing as the format in which they should be collected, nor does it mean identifying the sources of the information. This chapter offers suggestions about who might find the information and record it on your manual information system's data collection sheet or enter it in the computer. What you are really looking for is a set of skills and people who possess them.

Data Collection Skills

At this point, it is important to differentiate between the routine information-gathering activities the jail will perform to maintain a modest but effective manual or automated information system and the special data collections that are done for specific purposes (e.g., planning for a new facility or reducing crowding in the jail).

Let's start with maintaining the management information system. Assume that the data elements to be routinely captured have already been determined and that the mechanisms for recording this information have been designed. (Chapter 4 provides guidelines for sheriffs and jail administrators who are interested.) But if the organization does not have the skills to maintain the system, there is no point in setting it up in the first place.

The skills associated with maintaining the information system are good clerical and/or basic computer skills, including:

- Ability to read and write legibly,
- Ability to do simple math (mostly adding and subtracting) with the assistance of a calculator (if not using a computer),
- Ability to organize work,
- Working knowledge of a spreadsheet and/or data base program (if using a computer), and
- Commitment to doing the work on a daily basis.

Once the system is set up, clerical personnel can maintain it.

Collecting data for a specific study requires a few more skills, particularly when the data collection involves a lot of data and information must be collected from more than one source. Still, these skills are routinely found in jails, other agencies of county government, criminal justice planning agencies, consulting firms, and the community. The skills you will need are described next.

1. Data Collecting and Coding Skills. Once again, these are mostly good clerical skills. It is very helpful if the collector or coder has some knowledge of jail operations and criminal justice system terminology. Data collectors and coders should also know what information is already

being collected in the jail. They should be resourceful people who are not afraid to ask questions about the information they uncover. They will need access to people within the system who can answer their questions. Depending on the size, scope, and time constraints of the data collection, this person will probably need to work at least half-time on the data collection while tracking down information.

2. **Coordinating Skills.** These are mostly administrative skills, enhanced by some technical expertise. The data collection coordinator needs to have an understanding of basic descriptive statistics, good writing skills, some experience in using data to convey information and in making presentations, and some understanding of the jail. This person will most likely present the information. **He/she** will help make some critical data collection decisions — like how big the sample will be, what data elements will be included, and which sources of information will be used.
3. **Analytical Skills.** These are largely interpretive skills. The person (or preferably persons) involved in analyzing the information should be able to think logically and see the whole situation in terms of the relevant components. These people will draw on their experience and knowledge of the jail and the criminal justice system. They will try to answer the question, "What does this data mean to us?"
4. **Data Processing Skills.** These skills could vary a great deal depending on the nature of the equipment being used for the data collection. At a minimum, data entry skills (typing information into a computer terminal) will be necessary. In the case of the Inmate Profile Data Collection Forms included in this document, the services of a person skilled in use of spreadsheet or data base programs would be very helpful.

Regardless of the software used, sheriffs and jail administrators should work closely with the person who will enter and process the data when designing the data collection sheet or the computer forms. This is to make sure that:

- The data can be input as easily as possible and
- The computer program being used can process the data collected

Few things are more frustrating than to collect data you really want to use only to discover that the computer cannot handle it the way you collected it! This is rare, since most common statistical programs can process any data that is numerically coded, but it **can** happen. Make sure it does not happen to your data collection!

For those who are reluctant to try their hand at a computer, or who do not have access to one, this document includes a manual data collection system in Appendix C. However, it is much faster and easier to use a computer to process data (particularly for a large number of cases) than it is to calculate averages on an adding machine and use tally sheets to figure out frequency distributions. When this manual was first published in 1982, few sheriffs or jail administrators had personal computers powerful enough and with enough memory to do major data collections. Few of the automated systems allowed someone other than a computer programmer to examine and analyze data. Now, most sheriffs and jail administrators have access to computer hardware and software capable of handling even large data collections. However, if you do not have a computer, this chapter will give you some ideas about where to find one for free or for a low cost to pay for the amount of time you actually use it.

Where to Find the Skills You Need

Each jail, criminal justice system, county, and community is unique. As a result, it is not possible to tell sheriffs and jail administrators specifically where they can find people with the necessary skills in their communities. A few lucky ones may be able to find all the skills in one person, but most will find that two or more people are needed. While it may be easier to coordinate the data collection if only one person is doing the multiple tasks, a team approach to the data collection has distinct advantages. A number of strategies can be pursued to locate people with the skills needed.

1. **Community Volunteers**. Many communities have volunteer offices that coordinate placement of people who want to volunteer their skills for public service. This is a source of clerical help that cannot be overlooked by jails, particularly those with little or no clerical help of their own. If the county does not have a volunteer office, sheriffs and jail administrators should consider recruiting volunteers on their own. Churches often prove to be good sources of volunteer help. This chapter offers some suggestions to make working with volunteers in jails easier for everyone!
2. **Law Enforcement Reserves**. While most **Reserves** are interested in the more exciting aspects of working for the Sheriffs Department, they **are** a potential source of help for some special projects.
3. **Student Volunteers**. Many community colleges have criminal justice course work, and students in these programs are ideal volunteers. Not only do they work for nothing, but they also are usually very motivated because they will be earning credits or completing a required internship, and their professor serves as an outside supervisor who can help ensure that they cooperate and participate fully. You also get the added benefit of pre-screening individuals who may apply for jobs in the near future. The same colleges often have computer centers and computer classes, with students who need practical experience. A number of counties have found students to enter all the data and a professor or graduate student with the skills to complete the computer run to process the data. The only thing these counties had to pay for was the computer time, which, in several cases, cost the county less than \$200.
4. **Temporary Clerical Help**. While temporary help does cost money, most temps do not cost as much as a regular employee and may be needed for a relatively short period of time (e.g., 3 to 4 weeks).
5. **Other County or City Agencies**. Other county agencies may be willing to lend clerical staff to your department for a short period of time, particularly if they have an interest in the outcome of the data collection. This may be very beneficial in the criminal justice system, since data collectors and coders who are familiar with the court or prosecutor's record systems may collect data there more efficiently than individuals who do not know the "ins and outs" of the system. If the county has a data processing department, assisting the Sheriffs Department or the jail with this kind of an effort may be part of its mission.
6. **Clerical Staff**. One of the most underused resources in criminal justice agencies is clerical personnel. Not only do they have the necessary skills, but they also have first-hand working knowledge of the criminal justice system, and they know where to find information. Training the

clerical staff to collect and code data may have many benefits, both to the organization and to the personnel involved, especially when collecting data outside of your own agency.

7. **Department Personnel with Technical Skills.** More and more people in all professions, including in the sheriffs department, are computer literate and enjoy activities like "surfing the Net" in their spare time. Most people graduating from high school and junior college have more than a passing acquaintance with computer technology, and many have been introduced to statistics as well. Although these activities may not be in their job description, except in the phrase "other duties as assigned," they can be a valuable resource. You will need to decide whether this should become part of their official job or done on an overtime basis.

Working with People from Outside the System

Many criminal justice organizations use volunteers from outside the criminal justice system to help in special data collection efforts, particularly those associated with planning new jails. Citizens' advisory boards often take an active role in this process. While this manual cannot provide all the information needed to successfully initiate a volunteer program, the National Institute of Corrections Information Center has a great deal of information that could help you get started. In the meantime, here are a few tips that may help.

- Speak English, Not Jail. Criminal justice professionals have a language of their own. It is filled with abbreviations that are routinely used by people in the system, but are unfamiliar to outsiders. If you hope to collect valid information, make sure that everyone understands what you are talking about.
- Find Out Who You Are Dealing With. Set up a mechanism to screen volunteers. Most criminal justice organizations that use volunteers conduct the same background investigation on them as for employees. This is not to suggest that you should require psychological evaluations of volunteers, but routine reference checks and clearances should be part of the volunteer selection process.
- Provide Orientation and Training. Most people from outside the criminal justice system find jails very intimidating places to be. Volunteers will need an orientation to the jail, as well as a place to work in which they feel safe and relatively comfortable. Training must be provided in the tasks that the volunteers are to perform. For example, if volunteers collect and code data, training in good coding procedures and the codes being used is mandatory. If they will use the computer, training in the appropriate procedures is needed. If volunteers are going to work in the jail itself, they will need an explanation of the security requirements that apply to **them** and a clear explanation of the behavior expected of them while in the jail. They should also be clear about unacceptable behavior.
- Check Your Confidentiality Requirements. Before recruiting volunteers, check your State's statutes regarding the confidentiality of inmate records. If there is any question regarding the use of volunteers, an opinion from the County Attorney may be necessary. Generally, the laws regarding the confidentiality of records allow individuals to gather information from them for research purposes. To ensure the confidentiality of the records once the information is gathered, remove any identifying information, such as a name, from the data collection sheet. When in doubt, use caution and good sense.

- **Accept the Volunteers.** Law enforcement and correctional agencies sometimes find it very hard to accept "outsiders." This is not very conducive to developing good working relationships with volunteer groups. Volunteers will be doing you and your jail a favor, and being supportive of them is critical.

Conclusion

Doing a good data collection in-house is hard work. It is an area in which counties frequently seek outside help. And there is nothing wrong with securing professional help in this area. Sometimes it may be much more practical to seek outside assistance, as when the jail lacks the necessary skills or when there is not enough time to do the study in-house.

However, there are a number of benefits to doing the data collection in-house:

- Since you understand the criminal justice system better than outsiders, you may be able to identify better sources of information than they could. You may have access to information that outsiders do not and will never be privy to. You will not need help to understand either the language of corrections or the peculiarities of your organization.
- Your familiarity with the criminal justice system may let you make better "educated guesses" that shape the research.
- You understand the politics of your organization better than an outsider and will probably know more accurately the alternatives that will be acceptable to the criminal justice system.

Outside consultants cost money

Most important, if you do your own data collection, you are far more likely to believe the findings and, as a result, it is more likely that some action will be taken.

As you might expect, there are also drawbacks to doing your own data collection,

- Doing a data collection involves an investment of time and energy. And data collections are something that should be done right if you want to be sure that the results are accurate.
- Because you are so close to the situation, your interpretation of the data may be somewhat slanted. Having an outsider participate in the analysis and interpretation of the information can guard against bias.
- If you are not familiar with what is happening in similar jurisdictions, it may be difficult to identify data that suggests something unusual is happening in your criminal justice system. To illustrate, if you had little knowledge of other jail populations, you might not think it unusual that 40% of your jail population was female. An outsider who has seen data from other jails can help guard against this problem.

In deciding whether you are going to do your own data collection, you must realistically determine if you have the people with the right skills and the time to do the work. Chapter 4 identifies places to locate the necessary information and tips on useful data collection formats.

Chapter 4: How to Locate and Capture Information

In the past, people often began data collections by grabbing a pad and pencil, starting a tally sheet, and rummaging through files. Six pages later, unable to read their notes, they designed a form on which the data could be recorded. Halfway through the data collection, they discovered that they did not leave enough spaces for one of the data elements. At this point, they often changed their minds about doing the data collection.

On today's information highway, people begin data collections by turning on a computer, bringing up their favorite data base or spreadsheet program, entering some data element names, and rummaging through either computer or paper files. Halfway through the data collection, they discover that they did not leave enough spaces for one of the data elements. No problem—just insert a new column or field—and re-rummage through the files. And they cannot remember how they coded "Breaking and Entering" yesterday. Small problem—review all the data. At last, when they attempt their first sort, they discover—oops!—not enough memory. After trying to recode all the alpha characters as numeric; the data collection is abandoned.

This chapter suggests ways to avoid these difficulties by:

Identifying documents, forms, and logs commonly found in local jails, which are frequent sources for the data elements identified earlier;

- Suggesting how to implement an automated or manual information system, providing samples of the formats or forms that might be used in such a system, and estimating the time involved in maintaining it;
- Discussing guidelines for the development of data collection formats or sheets and code books.

Jail Information Sources

The first step in locating information is to identify all the places where it is already being captured or recorded on a routine basis. One way to do this is to take blank copies of every form routinely used and list the data elements that are specified on them. Next, review all the activity logs (such as the Booking Log, Release Log, Daily Activity Log or Shift Activity Sheet, etc.) and do the same. Do not list items that are normally recorded in any narrative sections, because:

- They will probably not be recorded on all the forms and certainly will not be found in the same place on the forms, and
Reading the entire narrative to extract one piece of information is far too time-consuming for most data collections.

Figure 4.1 is a traditional Incident Report form. The data elements that officers must fill out on this form are:

- Inmate's name:
 - Inmate's hooking number.
 - Inmate's housing assignment at time of occurrence:
Date of incident.
 - Time of incident:
Date report is completed. and
 - Time report is completed.

Reviewing the forms will almost certainly reveal that a number of data elements are already being collected. However, some essential data elements may be camouflaged in the narrative section of the form. If so, make a list of these items and consider revising the form so that it specifically asks for the essential missing items. For example, in the case of the Incident Report form in Figure 4.1, a data element that is missing and highly desirable is Location of Incident. While this is normally included in the narrative section, it is often not very visible. Sometimes officers forget to record the location or list several locations. This data element should be placed on the list of missing data elements. Later examples of this form suggest several ways in which the form could be modified.

Figure 4.1 Sample Incident Report

Any County Sheriffs Department
Jail Division Incident Report

It may be that a number of common data elements are not collected in your facility. If few forms are used and the only log book is the Booking Log, your jail does not have a very developed recordkeeping system. While that is beyond the scope of this document, it should be a matter of great concern to sheriffs and jail administrators. With the prevalence of litigation against jails and prisons, the sheriff, jail administrator, county commissioners: and the county itself are in a very vulnerable position without adequate documentation.

The logs and forms described below **capture** the data elements commonly collected in jails and provide documentation in the event of litigation.

1. **A Booking Log** is a chronological record of all persons who are arrested and booked at the facility. It typically captures:

- Date.
- Time prisoner accepted,
- Arresting agency and officer:
- Name of prisoner,
- Date of birth,
- Place of birth:
- Charges,
- Court of jurisdiction,
- Bond for each charge, and
- Booking number.

Do not be surprised if the Booking Log used in your jail is not exactly like this. The information recorded in the Booking Log will **vary** from county to county, but it should contain most of these data elements. The Booking Log is an important source document for both the Inmate Population and Inmate Profile Data Elements.

2. **A Release Log** is a chronological record of all individuals who are released from the facility. It typically includes:

- Date and time of release,
- Name of prisoner?
- Booking number, and
- Means of release (including information about bond).

Release Logs are often combined with Booking Logs so that release information is recorded on the same line as each inmate's booking information.

3. **An Arrest Report** provides basic identifying and some demographic information about persons who have been arrested, a statement of the authority under which they are detained, and some **information** about the details of the arrest. Jails usually keep a copy of the Arrest Report (or a comparable summary) as their documentation of the reason for the arrest. This practice enables them to prove that the jail is not unlawfully detaining anyone. The Arrest Report is a good source of some Inmate Profile Data Elements.

A word of caution is in order. Although there is some evidence that information that comes directly from people without any additional documentation ("self-report data" in research terminology) is quite reliable, it is important to note that Arresting Officers get information at a time when prisoners have good reasons to conceal the truth and make themselves "look good." As a result, data elements like employment status or education level may not be accurate.

4. A Bond Log is a chronological record of all bonds that are paid at the jail. It usually includes:

- Name,
- Date,
- Time,
- Charge (and bond for each),
- Type and amount of bond,
- Return **court** date,
- Jurisdiction, and
- Officer's initials.

5. A Medical Screening Form records the results of the Intake Medical Screening. It usually includes:

- **The** inmate's responses to a series of questions designed to reveal the presence of any medical conditions requiring immediate treatment or action by jail staff;
- A **brief** description of the inmate's physical state; and
- A brief medical history, including questions about substance abuse, mental health problems, and any treatment the inmate was receiving or had received in the past.

This form is one of the few places to find Inmate Profile Data Elements relating to medical, mental health, and substance abuse problems.

6. A Classification Screening Form records much information about the prisoner, including:

- Demographic information;
- Criminal history;
- Current legal status;
- Medical, psychological, and substance abuse problems, and treatment recommendations; and
- Program and service needs.

Classification Screening Forms may be relatively simple **forms** or very detailed documents resulting **from** an in-depth interview. They are excellent sources of Inmate **Profile** Data Elements.

7. A Daily Activity Log (or Shift Activity Sheet) is a chronological record of all operational activities. Whether the jail uses a Daily Activity Log or Shift Activity Sheet will depend on the size of the facility. Small facilities (fewer than 50 beds) will probably use a Daily Activity Log; larger facilities will probably use a Shift Activity Sheet. Very large facilities (100+ beds) will probably need to subdivide the data elements listed below into separate logs:

Court activity (who went to court for what **type** of hearing);

- Master control activity (results of routine head counts: inmate movement records: results of any routine tests done on jail equipment. notation of any unusual events—e.g., smoke detectors going off. etc.);
- Physical plant checks (routine checks of all locks, windows, security equipment. life safety equipment. etc.);
- Laundry exchange (dates of distribution of bedding and/or uniforms to prisoners);
- Outside agency inspections (dates and results of routine inspections by state inspectors, fire marshals, public health officers, etc.); and
- Program activity records (date, time, type of program—e.g., recreation: religious services. etc—and identity/number of inmates attending).

Other than the head counts, which would provide Average Daily Population, this information does not relate directly to any of the data sets identified in this manual. There is a great deal of variation in what these logs are like and how they are maintained. Some jurisdictions have automated this log in a way that allows them to extract information in summary form for management decisionmaking purposes. That is an ideal situation. Others have automated the information using word processing software, which makes it difficult to analyze without having a person extract and review the information. Others continue to use a manual log, which has the same drawbacks as a log in a word processing document. In both word processing and narrative formats, information is often inadvertently omitted or not recorded in the same way. These problems make logs of this type difficult to use for data collection and analysis.

8. A Visitor's Log is a chronological record of visitors to the facility. Both professional and family visitors should be entered in the Visitor's Log. Data elements found here include:

- Name of visitor,
- Name of inmate visited,
- Relationship to visitor.
- Date, and
- Time.

This is a good source of information about the level of activities in the facility, and its review is a critical component of a staffing analysis.

9. A Program or Activity Log is a chronological record of each type of program or service provided in the facility. These logs are often kept for activities that are required by standards (to document compliance), that are associated with grant funding (such as some drug and alcohol programs), or for budget purposes. Data elements found here include:

- Name or title of program.
- Number of inmates participating.
- Name or other identifier of service provider,
- Date, and
- Time or duration of program.

Like the Visitor's Log, these logs are good measures of activity in the facility. They are useful for budgetary purposes, defense against possible litigation, program development and staffing analysis.

10. A Transport Log is a chronological record of prisoner transportation activities. It includes all the data elements identified in the Transport Set mentioned in Chapter 2. It is often the only source of that information and is critical for staffing analysis. Facilities that do many transports may want to develop a Transport Form to be completed for each transport and then compiled in a monthly report. It should be possible to analyze a specific set of transports related to courts.

I I. An Incident Log or Incident Report Form is a chronological record or file of all "out-of-the-ordinary events" in the facility, including fires: assaults: injuries, **contraband**, major and minor rule infractions: suicide attempts, violent behavior, psychiatric problems, etc. It should include all the data elements noted in the Incident Data Set described in Chapter 2. This is often the only source of information about these events.

12. Personnel Records document all personnel actions taken, including:

- Hiring (date, salary, position):
- Terminations (date, position, type: reason): and
- Demographic information.

These are normally kept in individual employees' files; they are an important source of information for the Personnel Set. One way to facilitate using these documents (without having to go through every file) is to develop a Personnel Notebook, with a face sheet for each employee. Copies of all personnel action forms can be placed in the notebook behind each employee's face sheet with a special section of the notebook for the sheets of all employees who are no longer with the agency.

13. Budget Records document budget allocations and expenditures. The budget should be divided into categories so that costs attached to personnel and salaries, fringe benefits, food: medical care, inmate supplies: etc., can be used to determine jail costs per inmate.

14. A Jail List or Roster is a daily list of

All inmates detained at the jail: including their charge status, legal status: specific charge, **any** holds, housing location, gender, and age;

- All bookings; and
- Releases and transfers (means of release; type and amount if a bond).

This list is primarily put together for others in the criminal justice system and should be sent to all courts, the public defender, the prosecutor, law enforcement agencies, etc. It is a source of both Inmate Population Data and Inmate Profile Data. Maintaining this document is not as **time-consuming** as might be imagined, particularly considering data base (or even word processing) software capabilities. For the unautomated jail, a little creative use of correction fluid and the copy machine makes it possible to type only one new copy a week.

The documents listed above are found in most jails. The forms and titles may vary, but you should be able to identify them by their function. If you have questions about recordkeeping systems, the National Institute of Corrections (NIC) Information Center (1-800-877-1461) has additional resources that **might** help you.

In terms of data collection efforts, however: many of the forms and logs routinely kept by the jail are not particularly helpful. The reliance on narrative to record information is too time consuming and often results in omissions. The most useful information sources for data collections tend to be:

- The Arrest Report
- The Booking Log.
- The Incident Log or Report Form.
- The Transport Log or Form.
- The Classification Screening Form. and
- The Jail List or Roster.

The Problem with Data Sources— and Some Solutions

The documents kept in the jail provide little information about overall performance of the criminal justice system. It is possible to find out something about arrest, bonding practices: and sentencing for which jail time is given, but the whole picture is not available. As the jail data collection begins, some problems or quirks will emerge, because no system is perfect. Common problems are discussed below, but because each jail and criminal justice system is unique, a particular problem may not be included. However, the problems listed here are shared by virtually every criminal justice system. Some of them can be managed; others can severely damage the integrity of the data collection and should be avoided at all costs.

Problem 1: Data elements are scattered.

Typically, when jail data is being collected, the data elements are scattered among information sources. Some elements are available only outside the jail. Not only does this increase the amount of time required to collect the data but it sets other problems in motion.

Even within one criminal justice system, the same term can have different meanings. For example, the term "charge" has distinctly different meanings to law enforcement, the prosecuting attorney, and the courts. To avoid comparing "apples and oranges," sheriffs and jail administrators who use more than one source of information must make sure that everyone uses the same definition for the data elements they are collecting. Errors of this type can lead to mistaken conclusions in spite of the most sophisticated computer programs and statistical analysis.

Problem 2: Data is missing.

Missing data is a common problem and one of the most frustrating. When the collecting and coding persons finally find the right Arrest Report in the thick file on one of the jail's regular prisoners, they may discover that the one piece of information they are looking for is missing.

Missing data is frustrating from another perspective. Let's assume that a good sample of 500 prisoners has been developed for an Inmate Profile Data Collection. Unfortunately, because a lot of the data is missing, only 50 of the files have any information about one of the variables (e.g., employment status). Assuming that the information gathered about the 50 cases applies equally well to the other 450 people in the sample would be a serious mistake. Researchers consider this an "external validity" problem. While information is available about some of the prisoners, no evidence suggests that the results can be generalized to the other prisoners.

In fact there may be good reason to believe that the 50 prisoners who provided that information might be different from the other 450.

- Maybe they were unemployed people charged with serious offenses who stayed in jail a long time.
- Maybe they were all students who were trying to convince the Arresting Officer to take pity on them or
- Maybe they had good employment histories and were trying to make themselves look like good candidates for release on bond.

In those three "maybes" are three significantly different pictures of the employment variable. A little missing data is manageable in a data collection, but a lot of missing data makes it very difficult to interpret the results of the data collection and analysis. Any more than 10% missing data is a real problem.

However, there is some consolation. Missing data provides an extremely useful piece of management information. It means that staff are not filling out the forms completely. There may be a number of reasons for that:

- They do not have the information when filling out the form
- They forget to ask for the information.
- They forget to write down the information.
They do not think the information is important.

As a manager, you can excuse the first reason as long as the form is changed to reflect the information that is available when staff complete it. Then make arrangements to routinely capture the information when it does become available. However, you cannot consistently excuse the other reasons. They imply one of two things: Either your staff has to learn the importance of filling out the information or you need to design better forms. Two more implications are hidden in that statement:

- Staff should not be asked to fill out information that is not important.
- The information collected should be shared with staff and used to solve organizational problems.

Problem 3: Information on the form is wrong.

"Bad data" is a serious problem for any data collection. It is especially serious if you do not know that the information being gathered is wrong. Earlier in this chapter, we mentioned a problem associated with taking the Employment Status data element from an Arrest Report (i.e., the prisoner might have reasons to give inaccurate information). That problem belongs in this general category. One way to deal with this situation is to check the information given by the prisoner at one time with information given later or against information that has been verified.

Another kind of error should be mentioned. The people who fill out forms are human; they make clerical and typographical errors in recording information. While these errors may be obvious to you or to operational staff, they will be only a series of numbers to the person who enters data into the computer. Even worse, the computer will make no judgments at all about the quality of the information that it processes. One of the oldest computer programmer sayings...

Garbage In = Garbage Out

...is still **true** today

Sheriffs and jail administrators can take steps to avoid this problem by identifying a logical range of responses for each data element. For example, the likely range of responses for the Age variable might be from 18 to 55. Checking **anything** above or below that is a common procedure statisticians use to make sure they do not have bad data. Most computer printouts will list the high and low scores for each variable, which makes it easy to **perform** this simple check.

If you note responses that are significantly above or below expectations, the person who processed the data will know how to find out whether there is a glitch in the program or **problems** with the data itself. There are statistical conventions to help counteract the effects of bad data.

Problem 4: Data elements are poorly defined.

This can happen to the most skilled developers of data collection sheets. They design a question that asks the **person** filling out the **form** to make a value judgment and then fail to provide the criteria for making that decision. A common jail example follows.

Suppose the Medical Screening Form asks the Booking Officer to answer the following question: "Does the prisoner exhibit any mental health problems?" This sounds fine until you think about what constitutes a mental health problem. Does it include:

- *Only people who are so "out of it" that they have to be seen by a crisis worker from the Mental Health Department?*
- *People who are violent on Intake after a bad night at the local bar?*
- *People who smile benevolently and giggle at questions like "Name?" but are totally cooperative?*
- *Only people who would be considered legally insane?*

There is nothing **wrong** with asking this question, provided that the criteria for answering "Yes" or "No" are defined. Make sure to get some help in defining these criteria and then train the jail staff in applying them.

Problem 5: The handwriting is illegible.

Sometimes the data is there, but the handwriting is indecipherable to the data collectors and coders. And "guesses" can lead to bad data. This is something a manager can remedy for future data collections. If you have people on staff whose handwriting is not legible, suggest they print. If that fails, suggest they type. And if that fails, consider whether that person should be in that particular position.

Problem 6: No one keeps data for the whole system.

This is really an extension of Problem 1. Each of the agencies in the criminal justice system keeps its own data, and some guard that data well. As a result, each agency has facts about only its own part of the system, which results in information being lost or fragmented. It may be extremely difficult for a data collector from one part of the system to get information from or access to records in another part of the system. This can sometimes lead to different agencies maintaining duplicate record systems.

All of this lends credibility to charges that the criminal justice system is uncoordinated. As policymakers in the system, sheriffs and jail administrators must take an active role in remedying this problem. At a minimum: the elements of the system must share information and work together on common problems. But developing an automated, integrated information system for the entire criminal justice system may not be a feasible goal for many systems. It may be more feasible to develop an automated information system for the jail using off-the-shelf software.¹ Once the rest of the system sees the benefits of having management information on a regular basis, they might be more interested in working on a shared system.

Setting Up an Information System for the Jail

Information systems are structurally or functionally related elements or procedures that provide information to an organization, usually in summary form. They may be automated (computerized) or manual (on paper). They vary a great deal in their usefulness, their complexity, their “user-friendliness,” and the amount of time involved in maintaining them. The system described in this section is a very basic management information system that generates monthly statistical reports about the Inmate Population Data Elements and some Inmate Profile Data Elements, as well as some operational data. If you will not be automating the system: now would be a good time to turn to Appendix C.

Elements of the System

This system assumes that you have no other automated jail functions. If you do, your system should be providing you with much of this information already. However, even if you already have an automated system, do not skip this section of the manual. It will help you get some ideas about what your system should be doing for you. Automated information systems are discussed further in Chapter 9. To set up this management information system, the jail will need:

- A computer capable of running your selected off-the-shelf software;
 - Either spreadsheet software (such as Lotus 1-2-3, Quattro Pro, or Excel) or data base software (such as dBase, Paradox, Access, or Approach); and
- A regular source of the data elements selected.

System Start-UP

As discussed in Chapter 2, the first thing to be considered is to define the problem and associated data elements. A second critical decision is to determine which types of software will be used. Generally speaking, spreadsheets are very good at computing and calculating math, they were developed to add, subtract, multiply, and divide. Data bases are very good for storing and sorting information. While most spreadsheet software can do some data base functions, most spreadsheet software has limited ability to

*For those using a manual information system, refer to Appendix C..

sort data by a set of criteria. If you rely on a spreadsheet, you will need to do many sorts with many data ranges to get the information that a good data base program could easily provide with one query. On the other hand, data bases (particularly some older data bases and some that were specifically written for one user agency) find doing math awkward, even if they can sort by multiple criteria. As a result, there may be a tendency to use spreadsheet software for Inmate Population Data Elements, which record summary information like counts on a daily basis, and data base software for most other data collections, such as an Inmate Profile Data, especially large data collections.

It is also worth noting that most current off-the-shelf software has been written in a way that allows users to import or export data from one application to another. For example, you could store information in a data base, produce your frequency distributions in the data base program, export the data to a spreadsheet application to "crunch numbers," and then export your tables to a word processing program for inclusion in a report and to make the graphics. If you have an existing jail information system, you should be encouraged to know that it is almost always possible to retrieve the information that you have entered to perform a special analysis. The issue will be how easy or difficult it is to do that. Take the time to learn how to develop special reports, query the system, and export data.

Step 1: Identify the information the system should collect. It is hard to decide what information is routinely needed if you have not worked much with data and information systems. Sometimes sheriffs and jail administrators find it difficult to identify what they should be asking their systems to do for them. As a result, they may hesitate to ask. That puts them in the position of letting system designers decide what they need. Since most system designers do not have operational experience, they tend to miss a number of the Operational Data Elements. The purpose of the rudimentary information system in this manual is:

To collect the Inmate Population Data Elements (average daily population; length of stay; jail days; total bookings; net bookings; number of felons, misdemeanants, traffic offenders, and holds; number of pre-trial, post-trial, and sentenced offenders; number of male and female inmates; and adult and juvenile inmates):

- To collect some basic Inmate Profile Data Elements (charge, age, ethnicity, residence, and release status); and
- To collect some Operational Data Elements (number and location of transports; number of incidents; and number of contact, non-contact, professional, and family visits).

Step 2: Determine how frequently summary reports are needed. Most sheriffs and jail administrators will probably want monthly statistical reports. The Inmate Population Data Elements are most understandable statistically when grouped by month; weekly statistics are too much to absorb when looking for jail population trends and patterns. An exception occurs when the jail is trying to control crowding. Then monthly statistical reports would not be frequent enough; weekly summaries would be helpful.

Step 3: Design the format. Tips on how to design good data collection sheets and forms are given later in this chapter. Figure 4.2 provides a sample spreadsheet Inmate Information Format. A copy of the spreadsheet is available from the National Institute of Corrections. If you use a data base application to record this information, the entry screen can be designed to your specifications.

Figure 4.2 Sample Inmate Information Format

ID #	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay	Birth Date	Race	Sex	Employment	Charge	Residence	Release To	Bond Type	Amount

Step 4: Decide who will maintain the system. Sheriffs and jail administrators should make one person responsible for maintaining the information system. And remember, it is a clerical function. If there are no clerical staff in the jail: officers who work the graveyard shift or dispatchers might be able to maintain the system during slow periods. One of the benefits of having clerical staff maintain the information system is that they are not tied to a radio like dispatchers. They can leave their work area to get information if need be.

Step 5: Determine where the jail routinely captures information and decide how to get it from that source. It is hard to be very specific about this step because jails capture the information in so many places. It would be helpful to review the list of documents included in this chapter. The Daily Activity Log or a Shift Activity Sheet, Booking Log, and the Arrest Report will probably provide most of the Inmate Population and Inmate Profile Data Elements. The Operational Data Elements will come from the Transport, Incident, and Visitor Logs. Another potential resource for the Inmate Population and Inmate Profile Data Elements would be the status board found in most Booking Rooms or in Central Control that records the inmates who are in custody, some descriptive information, and their housing assignments. For jails that have jail or booking information systems, review each of the screens on which data is entered to identify the data fields needed for the analysis.

Let's assume, for this example, that the inmate information is already being collected on a Daily Activity Log. Figure 4.3 provides an example. Appendix C includes an 8½" x 11" copy.

Figure 4.3 Sample Daily Activity Log

Any County Sheriff's Department Jail Division Daily Activity Log	
IN:	OUT:
Time: _____	Time: _____
Name: _____	Name: _____
Charge: _____	Release to: _____
P.D.: _____	Bond: _____
Time: _____	Time: _____
Name: _____	Name: _____
Charge: _____	Release to: _____
P.D.: _____	Bond: _____
Statistical Summary: Beginning Count _____ End Count _____ #Felons _____ # Misdemeanants _____ # Traffic Offenders _____ # Other _____ # Pre-Trial _____ # Sentenced _____ # Females _____ # Juveniles (Male) (F e m a l e) _____	

Your task will be to automate the statistical **summary** on the bottom of the Daily Activity Log. This is an ideal spreadsheet **function**. Figure 4.4 provides an automated version of the Daily Activity Log's Monthly Statistical Summary in a spreadsheet format.

Figure 4.4 Sample Monthly Statistical Summary Format

Monthly Statistical Report:													
1	2	3	4	5	6	7	8	9	10	11	12	13	
Date	Begin Count	End Count	Bookings	Releases	No. Felons	No. Misde-meanants	No. Traffic	No. Pre-Trial	No. Sentenced	No. Holds	No. Males	No. Females	
1/1/96 ↓													
January 2 - January 30													
1/31/96													

By adding a few formulas to this spreadsheet, you can easily create your own internal check to guard against entry mistakes. The totals of columns 6 through 8, columns 9 through 11, and columns 12 and 13 should each equal column 4.

Let's also assume that the formats needed to collect the operating information about transports, incidents, and visitors from the logs are not available. For each, repeat steps 1 through 4 to decide:

- What information should be included in the format;
- How frequently the information will be needed;
- The design of the Transport, Incident, and Visitor Log Sheets; and
- Who will maintain the system.

Figures 4.5, 4.6, and 4.7 provide sample spreadsheet formats for Visitor, Transport, and Incident Logs; Appendix C provides full-sized forms.

Figure 4.5 Sample Visitor Log Format

Visitor Log							
Date	Time In	Time Out	Inmate Name	Visitor Name	Relationship	Contact	Noncontact

Figure 4.6 Sample Transport Log Format

Transport Log													
Date	No. Officers	No. on OT/Comp	Reason	No. Prisoners	Location 1	Location 2	Depart	Return	Meal Cost	Mileage Start	Mileage End	Vehicle No.	Gas \$

Figure 4.7 Sample Incident Log Format

Incident Log															
Inmate Name	Cell Assignment	Date	Time	Reporting Employee	Location	Reason	Rule Violation	CR?	Action	Restraint	Injury	Review Board	Review Board Date	Date Notice Served	Grievance

If you use a data base program, you will be able to design your own data entry screen. However, if the jail is not automated, before you can use these formats, you have to gather the information on the forms. The easiest way to get this information is to:

- Require that all incidents that occur in the facility be documented in the Incident Log. The sheriff or jail administrator must define for staff what kinds of events must be documented. He/she should review each form and then give it to the person maintaining the information system, to be placed in the Incident Log file.
- Require that all officers transporting inmates outside the facility begin a Transport Form prior to their departure from the jail and that they complete it as soon as they return. Have them give the completed form to the person maintaining the system, to be placed in the Transport Log file.
- Require that all persons who visit the facility are entered in the Visitor's Log. The forms should be placed in the Visitor Log file.

If the jail has more than one source for a particular data element, a decision must be made about which source should be used. Several criteria might help with that choice:

- One source may be more accurate than another, and
- The fewer sources needed to get the information, the better. When confronted by a choice between one source that has only one data element and another source that has more than one data element, unless there is a compelling reason to use the source with just one element, use the source that provides the most data elements.

Step 6: Train the person who will maintain the system in how to use it. Insist that it be done daily! Automated systems, which require staff to manually enter information into the computer, are very manageable if they are maintained on a daily basis. But if either manual or automated information systems are neglected, updating them when it is time to complete the monthly report is extremely tedious. Worse, if some of the information is missing, it may be impossible to get it once the inmate is gone and staff's memory has faded. Occasionally, it is also a good idea to make sure that the information entered in the Inmate Information Format is accurate. Before beginning, make sure that the staff person who will enter the information has been trained on the software. Most software has a tutorial that can get people started, and other training resources, such as adult education courses, can help with common software. You will also need to explain how each data element is to be entered. If you use codes (e.g., 1 = male, 2 = female), you should have a list of these codes available.

Step 7: Pick a day to start. Ideally, the information system should start on the first day of the month. It is wise to allow a few days of practice to be sure that the system is working properly. On Day 1, open the Monthly Statistical Summary, and make the first entry. Figures 4.8 and 4.9 show two sample spreadsheet formats with entries.

Figure 4.8 Sample Monthly Statistical Summary Format with Entries

Monthly Statistical Report												
Date	Begin Count	End Count	Bookings	Releases	No. Felons	No. Misde-meanants	No. Traffic	No. Pre-Trial	No. Post-Trial	No. Holds	No. Males	No. Females
1/1/96	67	73	8	2	44	19	4	33	26	8	60	7
1/2/96												
Average	67.00	73.00			44.00	19.00	4.00	33.00	26.00	8.00	60.00	7.00
Sum	67	73	8	2								

A review of Figure 4.8 shows what happens when you add formulas to the format. Counts are averaged, which will give the average daily population within the categories defined on the spreadsheet. Sums will provide total bookings and releases. If you add the counts rather than averaging them, the result is the statistic, Jail Days, discussed in Chapter 3. One final computer note: Be sure to save this computer file with a new name. That leaves the original file to use as a template, or model, each time you want to start a new monthly summary. It is a good idea to use a simple file-naming convention to identify the new file. Many people use the month and year (e.g., 0194SUM.WK3) to make it easy to identify the file. Use this file for the next month.

Figure 4.9 Sample Inmate Information Format with Entries

ID #	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay	Birth Date	Race	Sex	Charge	Residence
96-001	Doe, John	01/01/96	01:30	01/01/96	06:30	0.21	07101153	White	M	Criminal "Trespass	State
96-002	Smith, John	12114195	18:05	01101196	07:00	18.00	12114162	Black	M	DUI	County

Enter each inmate who was released on the date you selected. In the Figure 4.9 example, two inmates were released. Save this file with a new name (e.g., 0196INM.WK3) and save the original file as a template from which you can build later months in your information system. Use this file for the next month.

Figures 4.10 and 4.11 show examples of statistical summaries that might be developed from Visitor Logs. Completing this statistical summary involves counting the numbers of visits and visitors in each category. It can help document the number of visits that occurred during each visiting period. Put this together with information about the number of visiting areas that are available and it can be helpful in determining if the scheduled hours are adequate. However, Version 2 (Figure 4.11) of the Visitor Log Statistical Summary takes this process a little further and, in doing so, becomes more valuable to the jail operation.

Figure 4.10 Visitor Log Statistical Summary, Version 1

Visitor Log							
Date	Time In	Time Out	Inmate Name	Visitor Name	Relationship	Contact	Non-Contact
1/1/96	13:00	14:00	Jones, John	Jones, James	Brother	X	
1/1/96	13:30	13:45	Smith, Paul	Smith, Vanessa	Wife		X
1/1/96	14:15	15:00	Williams, Richard	Williams, Susan & Gregory	Mother, Son	X	
1/1/96	14:15	15:00	Anderson, Jeffrey	Richards, Marilyn	Friend		X

Figure 4.11 Visitor Log Statistical Summary, Version 2

Visitor Log								
Date	Time In	Time Out	Time	Inmate Name	Visitor Name	Relationship	Contact	Non-Contact
1/1/96	13:00	14:00	1	Jones, John	Jones, James	Brother	X	
1/1/96	13:30	13:45	0.25	Smith, Paul	Smith, Vanessa	Wife		X
1/1/96	14:15	15:00	0.75	Williams, Richard	Williams, Susan & Gregory	Mother, Son	X	
1/1/96	14:15	15:00	0.75	Anderson, Jeffrey	Richards, Marilyn	Friend		X
Number of Inmate Visits								4
Number of Visitors								5
Number of Contact Visits								2
Number of Non-Contact Visits								2
Hours of Visiting Provided								2.75
Average Hours per Visit								41.25 minutes

Version 2 takes advantage of the math capabilities of the spreadsheet program to calculate the number of hours of visiting provided. Think about how this information might be useful. It could:

- Document compliance with a standard that requires a specific amount of visiting;
- Be useful ammunition if litigation occurs around visiting issues;
- Together with staffing information, identify the impact of visiting in terms of staff hours involved; and
- Together with budget information, identify the cost of visiting.

Observation: Sheriffs and jail administrators typically collect statistical information like that shown in the Monthly Statistical Summary. Automated information systems, particularly booking systems, often collect information like that shown in the Inmate Information Format. However, even if jail operational functions, like visiting, are automated, the information contained in them is seldom summarized, and yet it is potentially some of the most useful in practical terms.

Review the other two operational formats provided, the Transport Log and the Incident Log, and automate those as part of your information system.

Step 8: Begin daily maintenance routine. Each day, make the entries on the appropriate format. Be sure to save and back up your work! The time that is needed will depend on your decisions about the information needed. A reasonable time target is 30 to 60 minutes per day.

Step 9: Prepare the first monthly report. One of the conveniences of automation is that it makes completing monthly reports almost automatic. As you make the last entries for the month in the Monthly Statistical Summary, and the Visitor, Transport, and Incident Logs in your information system, you are actually completing your Monthly Report. You might use some of the techniques identified in Chapter 8 to improve their appearance.

Summarizing the information from the Inmate Information Format is only slightly more complicated. Without access to a good data base program, you can use the data base properties of the spreadsheet to categorize and sort information.

Figure 4.12 shows a completed Inmate Information Format. This format could be a table in either your data base or your spreadsheet. To begin to summarize this information: use your spreadsheet's commands to identify the data range or your data base's table to be sorted. Then identify the variable? column, or field by which you want to sort. For example, you may want to sort by residence. In data base and spreadsheet terminology: residence is the primary key. After sorting, the result will resemble Figure 4.13. Most data bases and spreadsheets have another function that allows you to create tables. Continuing to explore "residence," look at the six different responses listed in Figure 4.13:

1. Unincorporated Boulder County,
2. City of Boulder.
3. City of Longmont
4. Other Municipality in Boulder County,
5. Out of State, and
6. Other Colorado County.

Assign a code number to each

1. Unincorporated Boulder County = (1):
2. City of Boulder = (2),
3. City of Longmont = (3),
4. Other Municipality in Boulder County = (4),
5. Out of State = (5), and
6. Other Colorado County = (6).

Complete the data distribution to develop the table shown in Figure 4.13

Each data element collected can be coded and analyzed. Some of the data elements, such as Length of Stay, can be averaged and may be very useful on a monthly basis. Other data elements may be more useful on an annual basis for either budgeting or for a section of the Department's annual report. This chapter provides more information about coding data elements later.

Step 10: Evaluate the information system in 6 months and make any necessary changes.
Every jail has its own peculiarities, and this system may not fit your facility. After 6 months, however, you should have enough experience with it to know what portions should be changed or adapted to fit your needs. You may want to add some data elements that are not included in the samples provided in this manual. Most spreadsheet and data base software allows you to do just that. Set a time to do this; and be sure to talk about it with the person who is maintaining the system. Evaluations that are not planned have a way of never happening. This information system when tailored to meet the needs of the jail, can be a powerful management tool. Take the time to improve on it!

Figure 4.12 Sample Inmate Information Format

ID#	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay in Days	Birth Date	Race	Sex	Charge	Residence	Release	Bond Amount	Type
96-001	Doe, John	01/01/96	01:30	01/01/96	06:30	0.21	07/01/53	White	M	Criminal Trespass	City of Boulder	Acme Bail Bonds	\$1,000	Surety
96-002	Smith, John	12/14/95	18:05	01/01/96	07:00	18.00	12/14/62	Black	M	DUI	City of Boulder	Self		
96-003	Henry, James	12/15/95	06:30	01/02/96	01:30	18.00	09/24/64	White	M	Probation Violation	Other City	Self		Cash \$500
96-004	Milton, Fred	12/30/95	07:00	01/04/96	18:05	5.00	07/06/66	White	M	DUS	Other City	Sentence Served		
96-005	Brown, Patricia	01/04/96	19:30	01/05/96	06:30	1.00	04/16/68	White	F	Simple Assault	City of Longmont	Self		OR \$500
96-006	Jones, Paul	12/19/95	18:05	01/06/96	07:00	18.00	01/26/70	White	M	Assault and Battery	City of Longmont	Sentence Served		
96-007	Black, Richard	01/09/96	15:54	01/10/96	01:30	1.00	11/07/71	White	M	Possession Controlled Substance	Boulder County	Self		OR \$2,500
96-008	White, Carolyn	01/10/96	03:41	01/11/96	18:05	1.00	10/03/68	White	F	Petty Theft	City of Longmont	Self		Cash \$500
96-009	Sanders, Robert	01/12/96	03:50	01/12/96	12:30	0.36	07/15/70	White	M	Resisting Arrest	City of Boulder	Self		Cash \$500
96-010	Johnson, Michael	12/20/95	15:41	01/13/96	18:05	14.00	04/25/72	White	M	DUI	Boulder County	Sentence Served		
96-011	Patterson, Ralph	01/05/96	01:30	01/15/96	01:30	10.00	02/20/72	Black	M	FTA	City of Boulder	Other agency		
96-012	Richards, Kevin	01/13/96	19:05	01/18/96	18:05	5.00	06/13/58	Black	M	Probation Violation	City of Boulder	State DOC		
96-013	McDonald, Alfred	01/18/96	12:30	01/19/96	01:30	1.00	03/24/60	Black	M	Hold	City of Longmont	Other agency		
96-014	Jefferson, Donald	01/11/96	18:45	01/20/96	18:05	9.00	04/29/38	Black	M	FTA	City of Longmont	Other agency		
96-015	Ray, William	01/22/96	16:30	01/22/96	21:30	0.21	06/04/52	White	M	DUI	Boulder County	Spouse		OR \$500
96-016	Larson, Oscar	01/23/96	14:25	01/23/96	18:05	0.16	03/16/54	Black	M	Possession Controlled Substance	City of Boulder	Self		Cash \$1,000
96-017	Ford, George	12/19/95	22:30	01/24/96	01:30	36.00	12/26/55	Black	M	Robbery	City of Boulder	Self		Property \$10,000
96-018	Ellis, Edward	01/24/96	13:05	01/27/96	18:05	3.00	07/02/60	White	M	DUI	Adams County	Sentence Served		
96-019	Kraft, Mark	01/27/96	00:30	01/28/96	01:30	1.00	04/13/62	White	M	Hold	Other City	Other agency		
96-020	Patton, Glona	01/24/96	21:05	01/29/96	18:05	5.00	01/23/64	White	F	Receiving Stolen Property	Los Angeles, California	Self		Cash \$2,500

Figure 4.13 Sample Sorted Inmate Information Data Element

Residence	Code	Number	Percentage
Unincorporated Boulder County	1	3	15%
City of Boulder	2	7	45%
City of Longmont	3	5	25%
Other Municipality in Boulder County	4	3	15%
Out of State	5	1	5%
Other Colorado County	6	1	5%
Total		20	100%

A word to those of you who may already have some degree of automation. Review the preceding section and ask yourself, “Does my system do this for me?” If it does, great! That’s one less thing you have to do. If it does not, you may need to determine if the system can do some of the things discussed. Begin by talking with your programmer or computer vendor to determine your system’s capabilities. For more on this topic, see Chapter 9.

Congratulations! You have just set up an automated management information system! You have a mechanism by which you can collect all the information you routinely need. Now, let's deal with those special issue data collections.

Special Issue Data Collections

Although your automated system may routinely collect a great deal of the information needed for jail management, sooner or later a special data collection will be needed. You may decide to:

- Build a new jail,
- Design a new inmate classification system or evaluate the current one,
- Evaluate inmate programs provided in the jail,
Change operational policies and procedures,
- Conduct an analysis to determine appropriate staffing levels in the jail, or
- Take a **proactive** approach to crowding.

To begin, some historical data may be needed. Depending on how the information is stored (manually or by automation) and how accessible it is, this may be as simple as downloading information from the jail's automated information system to a personal computer in a data base or spreadsheet format that allows you to manipulate the data without risking the integrity of the records. **Or**, it may be as tedious and time-consuming as searching through hard-copy records.

In either case, while the data available through the information system will be invaluable in these projects, additional information that the system does not provide will often be required. This section **describes** a process for designing these special data collection forms, including **tips** for designing easy-to-use forms. The appendixes provide copies of data collection sheets and code books for:

- Inmate Profile Data Collection (Appendix D),

- Incident Data Collection (Appendix E), and
- Transport Data Collection (Appendix F).

Designing Special Data Collection Forms and Code Books

The process for designing forms for special data collections is similar to the process for designing the format for the automated information system.

Step 1: Identify the information to be collected. If questions arise regarding this step, refer to earlier sections of this chapter.

Step 2: Determine where the jail routinely captures the information. Designing the forms for special data collections is similar to the process for designing the form for the manual information system. Furthermore, if good design procedures were followed to design the Log Sheets or forms on which you routinely collect information, this should be simple. The forms routinely used will double as data collection sheets. If the jail's record system does not capture the information at all, there are only two choices to remedy the situation:

1. Redesign a form that staff already completes at a time when the information needed is likely to be available. For example, the special data collection may be concerned with finding out how many people in the jail have been hospitalized for mental illness. The jail's Medical Screening Form already asks whether the prisoner exhibits signs of mental illness. This would be an ideal time to ask about hospitalization for mental illness; the Medical Screening Form would be a likely target for redesign. This strategy limits the proliferation of additional forms, and it does not increase officer paperwork significantly.
2. Design a form to collect the information. To use the previous example, if you did not have a Medical Screening Form, this would be the time to adopt one. It is important to design good forms that capture all the needed information and reduce the amount of duplication.

If the Log Sheets and other routinely used forms are well designed, they will double as data collection sheets with no modification. A good example of a form that doubles as a data collection sheet is the Transport Log. However, if the information needed is located in several different documents, a separate data collection sheet will be needed on which all the information for each case is recorded. An example of this kind of form is the Inmate Profile Data Collection Form included in Appendix D.

Step 3: Analyze the data elements. Once all the data elements are identified, the next step to developing a special data collection sheet is to think about what information will be recorded for each data element. This is another time that first-hand working knowledge of the jail and its operations will be invaluable. To explain this step, let's work with the Transport Log Form. Figure 4.14 provides a sample of a completed Transport Log Form.

This step identifies the different ways in which data elements are measured. On the Transport Log, the responses to some of the data elements easily fall into different categories, such as:

1. Reason for transport — The items in the checklist define the number of categories that are possible for this data element.
2. Location from and to — All the locations that are possible places to which prisoners would be transported define the categories possible for this data element.

3. Vehicle used—There are three possible categories for this data element

Figure 4.14 Sample Completed Transport Log Form

Any County Sheriff's Department Jail Division Transport Log			
Officer Name and Badge Number:			
1. <u>George B. Good</u>	<u>#1617</u>	Regular Hours <u> </u>	Comp/O.T. <input checked="" type="checkbox"/>
2. _____	_____	Regular Hours <u> </u>	Comp/O.T. <u> </u>
Reason for Transport:		Date of Transport: _____	
<input type="checkbox"/> Emergency Room Treatment		Transport to State Penal Institution	
<input type="checkbox"/> Scheduled Medical Appointment		Transport to State Hospital	
<input type="checkbox"/> Court Appearance		Transport to Treatment Facility	
<input type="checkbox"/> Warrant Pick-up		Court Ordered	
<input type="checkbox"/> Return to Other Jurisdiction		Other: _____	
Inmate Name:		From	To
1. <u>Frank E. Felon</u>		<u>Jail</u>	<u>DOC</u>
2. _____		_____	_____
3. _____		_____	_____
4. _____		_____	_____
Time of Departure: <u>09:00</u>		Meal Cost:	
Time of Return: <u>19:00</u>		1. \$ <u> </u>	<u>3.99</u>
		2. \$ <u> </u>	<u>3.99</u>
Vehicle:		Total \$ <u> </u>	
Starting Mileage: <u>32,675</u>		Vehicle Used: <u>CO-17</u>	
Ending Mileage: <u>33,115</u>		Mechanical Condition: <u>0 1 2 3 4 ✓ 5</u>	
Total Mileage: <u>440</u>		Cleanliness: <u>0 1 2 3 4 ✓ 5</u>	
Gas Cost: <u>\$21.80</u>		Gallons Purchased: <u>20</u>	
Related Transport Problems: <u>None</u>			

4. Badge number—The number of commissioned employees defines the number of possible categories in this data element
5. Date of transport—There are 365 possible categories for this data element (which would probably be grouped into months for a study).
6. Return and departure times—There are as many possible categories as there are minutes in the day (which would probably be grouped into hours for a study).

Other data elements have responses that fall into some kind of order from better to worse. The data elements like that on the Transport Log relate to:

- I. Mechanical condition—Asks the officer to rank the functionality of the transport vehicle from good to poor.
2. Cleanliness—Asks the officer to rank the cleanliness of the vehicle from good to poor.

The rest of the data elements have possible answers recorded in equal units. These data elements could have a great many possible answers, such as:

1. Regular and O.T. Hour;—Answers could vary from less than an hour to several days, but they will all be expressed in hours (or minutes that can be turned into parts of hours);
2. Total Time—Answers could vary from less than an hour to more than a day, but they will all be expressed in hours or parts of hours;
3. Total Mileage—Answers could vary **from** less than a mile to many miles, but they will all be expressed in miles;
4. Gas Cost—Answers will vary from \$00 to a specific dollar amount, but they will all be expressed in dollars and cents;
5. Number of Gallons Purchased—Answers will vary, but they will all be expressed in gallons;
6. Number of Gallons Required to Fill—Answers will **vary** from close to 0 to the entire capacity of the gas tank, but they will all be expressed in gallons; and
7. Meal Cost—Answers will vary from 0 to the total **amount** the Department authorizes for food on transports, but they will always be expressed in dollars and cents.

Why is so much time and space spent on describing the different kinds of data elements'? To statisticians, this elementary **information** is a big deal. These are the levels of measurement associated with all data elements. The first group of data elements (whose answers can be divided into a series of categories) are called nominal data elements. The second group (whose answers can be ranked in some kind of order) are called ordinal data elements. The third group of data elements (whose answers were expressed in equal units) are called **interval and/or ratio** data elements. Statisticians divide the **third** group into two subgroups, depending on whether 0 is a possible response. For the purposes of this manual, they can be treated as one group.

Levels of measurement are critical in calculating statistics, because some statistical procedures are valid only with data of certain levels of measurement. So, if the statistical procedure applied is not appropriate for the level of measurement **the** resulting analysis of the data may be meaningless. Chapter 6 spends more **time** on problems of this nature. However, this fact has one very obvious application for developing data collection sheets: Nominal data elements with a relatively short list of possible categories (like Reason for Transport) lend themselves to checklists that make the forms easier to fill out.

The final task in analyzing the data elements is to decide the level of detail to be reported. For example, for Total Mileage, it is necessary to decide if the answer must be written out to the tenth of a mile or if responses can be rounded **off**. Either write this instruction on the **sheet** or design the sheet in such a way that officers must write the answer in the correct **form**.

This is a critical task because these categories of **variables** will become the finest amount of **detail** that your data collection sheet will measure. It is always possible to regroup or redefine the categories into larger topic areas, but it is impossible to get finer detail once the information is coded. The data element Charge is an excellent illustration of this problem. If, for example, charges are coded as "Crimes Against Persons," "Traffic Offenses," etc., it would be impossible to find out how many inmates are charged with specific types of offenses or to investigate the characteristics of inmates charged with DUI, assault, **etc.** Take the time to make sure the information is gathered in the finest level of **detail** that could possibly be important to the jail.

Step 4: Design a draft data collection sheet or format. Once all the data elements have been defined and the ones requiring special treatment (e.g., checklists, decimal places, etc.) have been determined, design of the data collection sheet itself can begin. Data collection sheets are just paper forms or computer formats on which people record the information collected. Statisticians call them "Data Collection Instruments."

Good data collection sheets have the following characteristics:

- They reduce the amount of writing to a minimum by pre-printing the names of the data elements. They are clear, easy to understand, and easy to fill out.
- They group related data elements in sections so that the person filling out the sheet has a logical sequence to follow.
- They are as short as possible (preferably one sheet of paper) without being crowded.

The first data element on the data collection sheet (and in the code book) should be a Case Identification Number (a unique number assigned to only one case). Remember, a "case" in the statistical sense can be a person, event, or object about which information is gathered. This is important in jail data collections because:

- The same person may be booked at the jail on more than one occasion;
- The same "booking" may involve several incarcerations (e.g., an inmate sentenced to serve weekends); and
- The same "booking" may involve several charges.

Depending on what you are analyzing the "case" could be the person, the booking, or the charge.

Thinking in terms of "cases" is so important it is also addressed in the next chapter. But for now, either fill in the Case Identification Number (such as 001,002,003, etc., through the last case) on **each** of the data collection sheets before you start using them, or you can number the data collection sheets after the data has been collected but before it is entered into the computer. But do number them. Otherwise, there is no good way to check and correct **entry** errors, which, unfortunately, happen in both manual and automated information systems.

Begin by looking at the data collection sheets that are included in this manual. Think about the format of the data collection sheet. Once the results are **satisfactory**, type the form either on paper or on the computer (the Master Format). **As** nice as it may look, **this** will probably not be the **final** data collection sheet. It needs to be tested first. This will require a code book.

Step 5: Draft a code book. Code books are documents that translate the data element categories (which are usually written words) into a numeric coding system so they can be entered more rapidly and computers can process them more quickly and efficiently. Code books identify:

- The data elements,
 - The categories for each data element,
 - The codes for each data element, and
 - The name of the computer field or column location for each data element (if the data will be computerized).
-

It is generally good practice to have the data analyst help develop the code book. **He/she** will probably have suggestions about ways to format the data collection sheet that will make it easier to enter the information into the computer.

The following example **illustrates** how code hooks work. The categories (which statisticians call values) of the data element (variable) Resident Status are:

- Metro County resident
- Your state resident, and
- Non-resident.

It is not very efficient to type words into the **computer**, and some computer programs can handle only numerical data. To deal with this problem, people designing data collection sheets develop codes (numbers or short abbreviations like "M" for male and "F" for female) to represent the different categories (or values) of each of the different data elements (or variables). So, for the above example, the six values of Resident Status developed earlier in this chapter would become:

- 1 Unincorporated Boulder County = (1),
2. City of Boulder = (2),
3. City of Longmont = (3).
4. Other Municipality in Boulder County = (4),
5. Out of State = (5),
6. Other Colorado County = (6).
7. **Other** = (7),
8. Not applicable = (8), and
9. Unknown = (9).

The code for the data collection is in some respects like Morse code or the 10 code: it makes communication with the computer more efficient. Three major challenges lie ahead. The first **concerns** the Categorical Data Elements (which statisticians call nominal variables). It is critical to identify all of the possible categories (or as many as possible). This is another area in which knowledge of corrections and the criminal justice system is very helpful.

There are a number of ways to get started. First decide what level of detail is needed. For example, continuing with the data element Resident Status, decide if the responses should be rather general (like the values identified above) or more specific (the specific city or town in which the prisoner resides). For some data elements (**e.g.**, Charge Status), this will be very easy. The criminal code will **define** some of the categories for you (**e.g.**, felony, misdemeanor, etc.). For other data elements, it will be **much** more difficult.

Next, do a little more research to identify all the possible categories. It is often a good idea to ask an associate to review a final list of categories. For added insurance, include a category "Other," with a space to write in **information**. If it is possible that the people who will fill out the form may not have the information, add two categories "Not Applicable" and "Unknown." These will be a big help in finalizing the data collection sheet.

The second major challenge is particularly important if data base computer software will be used to process the data. It concerns the interval and ratio data elements, which are measured in some type of

numerical unit These are almost always expressed in numbers already. The challenge is to identify the number of digits that will be in the largest possible answer to that question. Meal Cost, one of the data elements on the Transport Log Form, is a good example. Assuming that both dollars and cents are necessary, the code book must identify how many spaces should be reserved for that data element. It could be three digits (\$9.99) or four digits (\$09.99).

If it is likely that the cost for a single meal would exceed \$9.99, then a four-digit code is necessary, or information about the more expensive meals will be lost. It is better to use an extra digit if the size of the largest response is uncertain or if the form will be in use for a long time.¹

The third and final challenge is to review every question to ensure that only one question at a time is asked.

Let's take an innocent-looking question, which might be found on a Medical Screening Form: "Has the prisoner been hospitalized for treatment of mental health or substance abuse problems?" This kind of question may sneak through the coding stages of the data collection without causing many problems. However, when it is time to interpret the information, there will be no way to identify how many of those who were hospitalized were treated for substance abuse problems or for mental health problems.

Review the sample code books in the appendixes and begin to develop your own

Step 6: Revise the data collection sheet or format. Based on any changes made while developing the code book, revise the data collection sheet or format. This time make sure that the space next to each data element reflects the exact number of digits that will be included in that code. So, for Meal Cost, a four-digit code, the entry on the data collection sheet would look like this:

Meal Cost .

Step 7: Test it. No matter how thorough the job of designing the data collection sheet has been, no one has ever designed a perfect form on the first try. It is important to find out how well the data collection sheet works. Statisticians call this process "pre-testing the survey instrument." It is a good idea to ask someone who works in the jail to fill out the form and give you feedback about how well it works before the form is finalized. When you are confident that the data collection sheet or format is right, decide on a time period for the pre-test.

As the data collection sheet is tested, problems will quickly surface. Information may suddenly appear that is not included in the coding system; data may not fit into the categories that seemed so clear. As decisions are made about how to handle each instance, write them down in the code book.

Step 8: Finalize the code book and the data collection sheet or format. When the test period is over, there are several items to consider. First, review the code book for changes made to respond to coding problems. Second, review each data collection sheet and look for places where testers have had to

*Note: This is a much greater issue with data collections that use statistical programs and some data base programs

use the "Other" category. If the same things are frequently written in, it is likely that one or more categories have been omitted. Third, look for data elements that have a lot of "Do Not Know" responses. This can mean a couple of things:

- Staff should be directed to find the missing **information** if it is **important**,
The data element should be removed from the data collection sheet if it is not important, or
- The question is confusing.

Most **often**, it is the latter. The only way to find out is to ask the people who completed the form. Once the nature of the problem has been determined, after a final check with the data analyst, make the final copies of the data collection sheet or format and the code book.

Conclusion

That is about all there is to it! The data collection sheet and the code book have **been** developed. Once the data collection sheet or format has been finalized and the data collection begun, your attitude toward changing it should be like **your** attitude toward drawing a gun: If there are compelling reasons, by all means go ahead, but if you can avoid it do so at all costs!

By keeping the data collection essentially the same from year to year, it is possible to compare the data elements over time. Statisticians call this a "longitudinal" data collection. If you do the data collection just once, statisticians call it a "cross-sectional" data collection. While a cross section can help you describe something, longitudinal data can document change and even begin to explain it.

So, without further delay, let's take care of a few final tips about **putting** the actual data collection together.

Chapter 5: How to Put It All Together

Now that the data collection sheet or format and the code book have been developed, only a few more details need to be cleared up before starting the data collection. Now is a good time to think about the details of really **doing** the data collection. That is what this chapter is about.

Doing the Data Collection

Doing the data collection is just a matter of good planning, organization, and some hard work. Collecting data in jails can be tedious and time-consuming. This manual has done its best, however, to make it as easy as possible to collect jail data by insisting that the proper preparations are made and by providing a number of guidelines and resources to save you extra time and energy. Now it is time to go to work. As you start the data collection, follow these steps to get organized.

Determine an Overall Strategy

First, determine where your jail routinely captures information and (if necessary) decide how to get it from there to a data collection sheet or format. This manual dedicated a considerable amount of time to this topic in the previous chapter. Documents that are common sources of jail information were identified. If you are operating the jail with a totally manual system, you will need to follow each step in the manual, using a handwritten data collection sheet (see Appendix C for instructions on doing a manual data collection). If you have automated your information system as suggested in the previous chapter, you may have some or all of the information you need already collected. If you already had an automated booking system, the information you need may be available there.

If your information is already automated, you still need to determine how to copy it to the file in which you will do the analysis. And you do want to copy it, to avoid the possibility of accidentally damaging or losing the data. Depending on your software, you may need to:

- Copy the files you want to analyze, or
- Extract the records you want from your data base.

This may involve translating the data from one computer application that stores your information to another that will do the statistical analysis.

Determine How Much Data to Collect

For some of the data elements (Inmate Population Data Elements and the Operational Data Elements), determining how much data to collect is not an issue. Information will be collected about every inmate or event. But for other types of data collections, determining how many and which cases to collect is a major issue. Statisticians refer to this process as sampling.

Although it may be desirable from a statistical point of view to have a large amount of data (e.g., Inmate Profile Data for an entire year), in reality, that is often not possible. Sometimes jail records have not captured the necessary information, or the record system may be unreliable. If the record system is in "good shape," historical cases can be used, but, in many situations, the data collection has to move forward in time to gather information about people who are currently passing through the jail. However, most systems cannot afford to wait for an entire year to get the information they need. Data collections need to be timely as well as accurate. Furthermore, collecting some kinds of data about every case may be too time-consuming. This is particularly true when the data collection involves getting data from other parts of the criminal justice system.

Note: Automated systems vary considerably in terms of what they do with the data when someone is released. In the worst case, the information is simply discarded or overwritten. In the best case, the most important information is retained in a master index file. Most automated systems ultimately have to deal with disk storage limitations. Old information may have to be backed up and removed from the disk. As a result, if you are using historical data, you may find that the information you need is not as accessible as you might wish. Ask your data processing department or computer vendor to help determine the best way to access the information you need.

Fortunately, statistical sampling allows analysts to gather information from a segment of the entire population (a statistical sample) in such a way that it is safe to apply what you have found in the sample to the whole population. You can even determine how safe you are in making this assumption. Earlier, this manual indicated that statistics enables analysts to estimate what the chances of being wrong are. Statistics uses the laws of probability and applies math formulas to what is known about the data. The laws of probability should be familiar to individuals who have ever placed a bet; the laws of probability are essentially the same thing as "the odds."

In sampling, statisticians use probability to determine what the odds are of selecting a statistically similar sample. Samples are statistically similar if they are drawn from the same population (all of the cases from which the sample is drawn) and the differences between them can be explained by the sampling process itself (sampling error). Statistics can determine how many times out of a hundred a sampling scheme will result in a statistically similar sample. Most statisticians want to be at least 95% certain that they would get similar results if they analyzed a different sample from the same population, and some want to be 99% certain.

The similarity of samples taken from the same population is measured by a statistic called the "standard error of the mean." Statisticians can also estimate the accuracy of this statistic, which indicates how reliable the sample is, as are the statistics that result from it. Researchers have to decide how big a difference they can tolerate in the results of their studies when they use a sample. This concept, "reliability," is meaningful when scientists are determining appropriate dosage levels for new medications; it is less a "life and death" issue for social research.

One of the main factors that causes the differences between samples drawn from the same population is something that statisticians call "variability." Variability examines how different the individual cases that make up the entire population are from one another. The more the cases differ, the more likely that samples drawn from the population will also differ. In other words, the statistical reliability will be less.

So, to construct a sample, you first need to answer three basic questions:

1. What is the population to be sampled?
2. How many cases should go into the sample?
3. Which cases should go into the sample?

In most jail data collections, defining the population is relatively easy. For example, a "typical population" would include "all persons booked at the Metro County Jail during calendar year 1997." To determine exactly how many cases should go into a sample, statisticians and mathematicians developed a formula. It is:

$$SE_x = \frac{s}{\sqrt{n}}$$

SE_x = The standard error of the mean

s = The standard deviation of the sample

n = The number of cases in the sample

This manual introduces "standard deviation" in the next chapter. For now, think about the concept rather than the specifics. This formula forces the user to decide two things:

1. How much variation can exist in the sample—in other words, how sure of the results do you want to be? 99% certain? 95% certain? or less? and
2. How reliable must the results be—how much error can be tolerated? This gets interesting when estimating dollar variables.

Most researchers do not follow this process every time they put together a good sample but use a table developed by statisticians to save time. The tables give the actual number of cases that must be included in the sample to meet the specifications determined by the answers to the two questions above. The tables included in Appendix G are for 95% and 99% confidence intervals. Variation in the sample error of the mean of 2%, 3%, and 5% are calculated for both the 95% and 99% confidence intervals. All you need to know is:

1. The level of confidence the analysts must have in the results (most jail data collections can live with 95% confidence intervals),
2. How reliable must the data be (most jails can get along with 2% or 3% sampling error, which is the same as 97% or 98% reliability),
3. The size of the population from which you are sampling, and
4. What proportion of the population from which the sample is drawn has the characteristic being measured (the tables include only proportions from 5% to 50% because if the proportion is greater, the question should be reversed). For example, if the Inmate Population Data shows that 20% of 1,000 annual admissions are women, the table will tell you that a sample of 198 cases should result in a sample in which 20% of the cases are women (see Table 5.1).

Table 5.1 Sample Size Example

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	64	108	165	196	213	217
1,000	68	121	198	244	262	278
1,500	70	126	211	269	297	306
2,000	70	130	219	278	312	322
2,500	71	131	224	286	322	333
10,000	72	136	240	313	356	370
50,000	72	137	245	321	367	380
100,000	73	138	246	322	369	384

The selection of sample size is particularly important when you are looking at a subset (segment) of the population. For example, pre-trial felons often have a disproportionate impact on the jail because of their long lengths of stay. Sometimes in the pre-trial process, felony charges may be dismissed or reduced prior to bind-over to the upper court. Using the percentage of arrests that are bound over to the higher court is an effective way to develop a statistically reliable sample size of pre-trial felony bookings.

Once the number of cases to go into the sample has been determined, the next problem is to decide which cases to include. This is important because the laws of probability (and all the statistics that are based on them) assume that the cases in the sample are selected randomly. This is a "random sample," which means that each case has an equal chance of being selected for the sample. Samples that are structured this way have the least bias (built-in error).

There are a number of ways to determine which cases go into the sample (sampling schemes). They include:

1. Simple Random Sample, in which cases are selected randomly from a homogeneous population.
2. Stratified Random Sample, in which the population is divided into groups, such as felons and misdemeanants, and cases are selected randomly from within those groups so that equal numbers are selected from each group. This kind of a sample would include an equal number of felons and misdemeanants.
3. Proportional Random Sample, in which the population is divided into groups and cases are selected randomly from within the groups until the groups are in the same proportion in the sample as they are in the population (e.g., if the population had 70% felons and 30% misdemeanants, so would the sample).
4. Cluster Sample, in which the population is clustered into specific areas or units (e.g., housing units). Then the housing units to be sampled would be selected randomly, and information collected about every individual in those housing units.

5. Cohort Sample is essentially the same as a cluster sample, but is based on time periods rather than physical or geographic areas (e.g., all prisoners arrested on a randomly selected date).
6. Systematic Sample, in which information is collected about cases in a certain sequence until the number of cases required for the sample is reached. For example, data is collected about every third, fifth, or tenth, etc., person booked (every "Nth" person in statistical terms) until enough cases are reached, with the number on which you begin being randomly selected.

Some of these sampling schemes seem to work better than others. Appendix H discusses methods for constructing a simple random sample and provides the random number tables (charts of numbers without any logical sequence) often used to develop random samples. Constructing random samples from random number tables can be extremely tedious and time-consuming. Statistically, they are "the best," but other sampling methods may be preferable.

Because of the nature of jail recordkeeping practices, a Systematic Random Sample, based on the booking or release log, works well for Inmate Profile Data Collections. Once the number of cases to be included in the sample has been determined, divide the number of cases in the population by the number of cases wanted in the sample. The result tells you how frequently to put a case in the sample (e.g., every third, fourth, or Nth case). For example, if the jail books 10,000 people per year and a 20% sample is wanted, information should be collected on every fifth case in the booking or release log. Do not start with the first case in the year. Randomly pick a number from (in this case) one to five and start with that case. This method will also work well if most of the information is gathered from records.

If possible, it is preferable to use a release log (rather than a booking log) as the basis of the sample. This helps to eliminate a particular type of bias with regard to Length of Stay data. Even over periods as long as a year, if information is collected on intake, at the end of the year some individuals in the sample may still be in custody, with no release data. In addressing this problem, the two most common approaches are either to leave Length of Stay as missing data (omit this data element on the coding sheet or format) or to use the last date of the period for which data is being collected as the release date (e.g., if the sample period were 1995, the release date would be 12131195). This biases the sample by making the Length of Stay shorter than it really is. This problem is particularly acute if data is being collected for only a short period of time.

Another type of sample that works well is a Cohort Sample. Randomly select dates throughout the year and collect information on all individuals in custody on those dates until the necessary sample size is reached. These samples are frequently called "snapshots" of the jail population.

Edward Lakner of the National Clearinghouse for Criminal Justice Planning and Architecture has developed a manual, *Statistical Sampling Methods for Correctional Planners*, that provides much information about sampling procedures. If determining sample size is difficult, perhaps a professional viewpoint will help. In general, samples smaller than 10% of the total population and less than 500 in absolute size make most statisticians uncomfortable. When in doubt, err in the direction of including too many people in the sample rather than too few.

Coding

The person who will collect and code the data must be thoroughly trained to fill out the **data** collection sheet or format. The **rules** that follow are good training points, but the data entry person also needs basic computer operation and software training. Many people underrate the importance of good data collection and coding practices, but this makes the difference between a data collection that provides

accurate information and one filled with errors. A number of coding practices will make the task easier for anyone who enters the data into the computer. The degree of automation and the type of software being used can make the coding process much easier.

- If a totally manual system is being used, you should follow the steps in the manual to code the data.
- If the system is automated, at least some of the data may already be coded.
If you are gathering data from multiple sources and are not part of an integrated criminal justice information system, you will probably need to use either a manual or automated data collection sheet.

Rules for Coding Data

The following are guidelines for coding data on the data collection sheet or format:

Rule 1: Use only one meaning for each code. Never use more than one meaning for a single code. For example, the Inmate Profile Data Collection Form and code book include several blank spaces to list city police agencies that bring prisoners to the local jail. Assign one city to each value of this variable (e.g., **6** = City of Metropolis). This is just as true for automated systems as it is for manual data collections.

Rule 2: Always write legibly. Be sure to write large and legibly if the data must be entered from a data collection sheet rather than directly into the computer. Data entry personnel who cannot read the information may enter the wrong codes, which will result in wrong analysis from the computer.

Rule 3: Limit the number of people who code the data. If possible, have only one person do all the coding to ensure that the information is coded consistently. If that is not possible, it is essential that all those who code information understand each data element and code it in exactly the same way. The person who coordinates the data collection should verify that this is happening.

Rule 4: Be consistent. Once codes have been established, do not change them! It will bias your information. If questions arise during the data collection, make a decision, record the answer in the Code Book, and make sure that all coders know how to handle this coding irregularity.

Rule 5: Make a copy of the last version of the code book. Without the code book, it is extremely difficult to analyze the data collection. If the same study is to be repeated at a later date for purposes of comparison, information must be collected in **exactly** the same format. Without the code book, that is next to impossible.

Rule 6: When in doubt, consider it as missing data. When in doubt about information to be coded (e.g., the handwriting is so bad that you cannot tell if the officer meant "Burglary" or "Battery"), treat it as missing data (leave the coding sheet blank or use your code for missing data).

Before the data collection begins, train the coders in good coding practices. Go over the code book item by item so that all the coders understand each one of the data elements. This avoids a situation such as having a coder ask something like, "Oh, you mean robbery is a crime against persons, not a property crime?"

Coding always goes slowly at first. There are questions to be answered, and coders need to become familiar with the code book and the information sources. Be patient. The more time people spend coding, the faster they get. The time spent up front to learn the coding scheme and to develop good coding habits is worth it!

Coding in an Automated System

This section of the manual is for those who **have** automated their information system as outlined in the previous chapter. Everyone else can **turn** to the next section, Gather Your Data Collection Supplies. If you are using spreadsheet software, you should be able to automate coding. If you are using one of the more "user friendly" data base programs, you should not need to actually code the data. Figure 5.1 shows a section of a typical Inmate Information Format. The advantage of both spreadsheets and data bases is that they generally allow you to enter information in alphabetic characters. However, to analyze the information, in the case of spreadsheets, you will need to do some coding, and in the case of data bases, you may have to "refresh" some of the information entered. For example, assume you want information about the charges on which an inmate was arrested. The following steps illustrate how to move from a list of charges to a table (or, in statistical terms, a frequency distribution).

Step 1: Define the data range. In some spreadsheet programs, you will define the data range by selecting the entire spreadsheet except for the column headings. In Figure 5.1, the data range consists of all the cells that are not shaded. In a data base program, select the field you want to sort by placing **your** cursor there. In Figure 5.1, each of the column headings represents a field.

Figure 5.1 Sample Inmate Information Format

ID #	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay in Days	Charge
96-001	Doe, John	01/01/96	01:30	01/01/96	06:30	0.21	Criminal Trespass
96-002	Smith, John	12/14/95	18:05	01/01/96	07:00	18.00	DUI
96-003	Henry, James	12/15/95	06:30	01/02/96	01:30	18.00	Probation Violation
96-004	Milton, Fred	12/30/95	07:00	01/04/96	18:05	5.00	DUS
96-005	Brown, Patricia	01/04/96	19:30	01/05/96	06:30	1.00	Simple Assault
96-006	Jones, Paul	12/19/95	18:05	01/06/96	07:00	18.00	Assault and Battery
96-007	Black, Richard	01/09/96	15:54	01/10/96	01:30	1.00	Possession Controlled Substance
96-008	White, Carolyn	01/10/96	03:41	01/11/96	18:05	1.00	Petty Theft
96-009	Sanders, Robert	01/12/96	03:50	01/12/96	12:30	0.36	Resisting Arrest
96-010	Johnson, Michael	12/30/95	15:41	01/13/96	18:05	14.00	DUI

Step 2: Sort by charge. In a spreadsheet, define the primary key for the sort and specify alphabetical order (ascending). If you are using a data base program, select the field (or fields) you want to sort, define whether you want ascending (A-Z) or descending (Z-A) order, and proceed with the sort. Figure 5.2 shows the result of a sort in ascending alphabetical order by charge.

Figure 5.2 Inmate Information Format Sorted by Charge

ID #	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay in Days	Charge
96-006	Jones, Paul	12/19/95	18:05	01/06/96	07:00	18.00	Assault and Battery
96-001	Doe, John	01/01/96	01:30	01/01/96	06:30	0.21	Criminal Trespass
96-002	Smith, John	12/14/95	18:05	01/01/96	07:00	18.00	DUI
96-010	Johnson, Michael	12/30/95	15:41	01/13/96	18:05	14.00	DUI
96-004	Milton, Fred	12/30/95	07:00	01/04/96	18:05	5.00	DUS
96-008	White, Carolyn	01/10/96	03:41	01/11/96	18:05	1.00	Petty Theft
96-007	Black, Richard	01/09/96	15:54	01/10/96	01:30	1.00	Possession Controlled Substance
96-003	Henry, James	12/15/95	06:30	01/02/96	01:30	18.00	Probation Violation
96-009	Sanders, Robert	01/12/96	03:50	01/12/96	12:30	0.36	Resisting Arrest
96-005	Brown, Patricia	01/04/96	19:30	01/05/96	06:30	1.00	Simple Assault

Step 3: Determine how often each charge appears. Examples do not usually include many cases. In the example, for instance, there are few enough cases to allow the responses in each type to be counted. However, in the real world of data collections, there will be too many cases to count efficiently. The software package you are using should be able to count the number of times each charge appears. Begin by reviewing the charges to make sure that the same charge is listed the same way each time. Note that the computer may treat capitals and small letters as different symbols and that simple assault, assault 3, and assault III will not be listed together. This is equally true of both data bases and spreadsheets. Once you review the charges that have been sorted, you may need to edit them so that they will be grouped together for easier coding and then resort them. When you are satisfied with the list of charges, assign a code to each unique charge. Figure 5.3 shows what the spreadsheet looks like when the codes have been added.

Figure 5.3 Sorted Inmate Information Format with Codes

ID #	Name	Book-In Date	Book-In Time	Book-Out Date	Book-Out Time	Length of Stay in Days	Charge	Code (1)
96-006	Jones, Paul	12/19/95	18:05	01/06/96	07:00	18.00	Assault and Battery	1
96-001	Doe, John	01/01/96	01:30	01/01/96	06:30	0.21	Criminal Trespass	2
96-002	Smith, John	12/14/95	18:05	01/01/96	07:00	18.00	DUI	3
96-010	Johnson, Michael	12/30/95	15:41	01/13/96	18:05	14.00	DUI	3
96-004	Milton, Fred	12/30/95	07:00	01/04/96	18:05	5.00	DUS	4
96-008	White, Carolyn	01/10/96	03:41	01/11/96	18:05	1.00	Petty Theft	5
96-007	Black, Richard	01/09/96	15:54	01/10/96	01:30	1.00	Possession Controlled Substance	6
96-003	Henry, James	12/15/95	06:30	01/02/96	01:30	18.00	Probation Violation	7
96-009	Sanders, Robert	01/12/96	03:50	01/12/96	12:30	0.36	Resisting Arrest	8
96-005	Brown, Patricia	01/04/96	19:30	01/05/96	06:30	1.00	Simple Assault	9

(1) =Values Range

After sorting the data, make sure that there are no variations within the same category (*i.e.*, the same charge is always written the same way). If you find duplicate entries, edit the data and correct this problem.

Step 4: Distribute the data. If you are using a spreadsheet, you can start to distribute the data by listing all of the codes on your spreadsheet as shown in Figure 5.4. Many data base programs will perform a count without having the data coded, which eliminates this step.

Refer to the manual for your software for the exact commands and steps it uses to distribute the data and perform mathematic calculations, such as sums, averages, counts, standard deviations, etc. Figure 5.5 shows the results of a spreadsheet with distributed charge data, and Figure 5.6 shows the completed version with percentages.

Figure 5.4 Table Prepared for Data Distribution

Charge	Code (1)	Number	Percentage
Assault and Battery	1		
Criminal Trespass	2		
DUI	3		
DUS	4		
Petty Theft	5		
Possession Controlled Substance	6		
Probation Violation	7		
Resisting Arrest	8		
Simple Assault	9		
Total			

(1) = Bin Kange

Figure 5.5 Distributed Charge Data

Charge	Code	Number	Percentage
Assault and Battery	1	1	
Criminal Trespass	2	1	
DUI	3	2	
DUS	4	1	
Petty Theft	5	1	
Possession Controlled Substance	6	1	
Probation Violation	7	1	
Resisting Arrest	8	1	
Simple Assault	9	1	
Total		10	

Figure 5.6 Completed Charge Table

Charge	Code	Number	Percentage
Assault and Battery	1	1	10%
Criminal Trespass	2	1	10%
DUI	3	2	20%
DUS	4	1	10%
Petty Theft	5	1	10%
Possession Controlled Substance	6	1	10%
Probation Violation	7	1	10%
Resisting Arrest	8	1	10%
Simple Assault	9	1	10%
Total		10	100%

Congratulations! Not only did you code all your data, but you also put it into a format that can be used in the document developed as a result of the data collection.

Gather Your Data Collection Supplies

Step 1: Copy the data collection sheets and code books. Make as many copies of the data collection sheet and code book as you will need to complete the data collection. Each coder should have a copy of the code book and several copies should be filed in reserve.

Step 2: Develop working files for the collection sheets. If this is a completely manual process, purchase several large file folders or manila envelopes to hold the data collection sheets. These will be useful if coders have to go to several sources for the information or if the information is collected over time. Label them "Completed Data Collection Sheets," "Blank Data Collection Sheets," and "Pending Information." If the system is automated, create the computer data entry **formats**. These may be a single file (for one coder and one machine) or multiple files (for multiple coders) that may be combined at a later time.

Begin

For some data collections (e.g., those that use historical data from already completed records), beginning the data collection will be a matter of pulling information out of paper files or from an existing data base. If the data collection is manual, the information is recorded on the data collection sheet, which is then placed in the "Completed Data Collection Sheets" folder. If the data collection is automated, the information is entered in the data collection format and saved when the information is entered.

Remember, always work with a copy of data from your automated information system. This protects **you** against unintentionally changing information in your records.

Other data collections, ones in which the information will be gathered over time or from multiple sources, are more complex. An Inmate Profile Data Collection, which collects information from a variety of sources about people currently in custody, is probably the most difficult logically. In that case, follow the steps below.

Step 1: Begin a data collection sheet or format for each person in the sample. Start a data collection sheet or format for each person selected for the sample.

Step 2: Organize them. Usually, the identifier on the coding sheet will be Inmate Name. In that case, file the data collection sheets in alphabetical order or sort the file alphabetically. Often Inmate Name is the **only** identifier that all of the elements of the system may have in common. If a numerical identifier (e.g., booking number) is used, file or sort the data collection sheets in numerical order.

Step 3: Secure information from the source documents. Go to the information source that provides the most information about each individual and complete as much as possible of the data collection sheet or format. Then go to the next information source. Continue until each information source has been exhausted and no more information is available about any of the cases.

Step 4: File partially completed forms. Place all the partially completed forms for persons who are still in custody in one file folder. If you are entering information into an automated format, you may elect to keep complete records in one file and incomplete records in another to conserve computer memory or to make it quicker to enter new information.

Step 5: Update release information daily. Each day, as people are released from the jail, pull their data collection sheets, add the release information, and review the forms to see if any additional information can be obtained. If so, file them in a "Pending Information" folder or keep them in your entry file. If not, place them in the "Completed Inmate Data Collection Sheet" folder or combine them with the new file that you created in Step 4.

Step 6: Add cases to the sample daily. Each day, as people are booked into the jail, begin a Data Collection Sheet for new bookings or add them to the entry file following the sampling scheme.

Step 7: Stop sampling when the necessary number of cases is reached. When the number of cases required for the sample are in the "Completed Inmate Data Collection Sheet" folder, the data collection is done.

Automation Issues

Several problems occur when you automate a data collection that involves multiple sources. The most significant is how to appropriately match cases. The criminal justice system tends to track people, events, and charges. One person can be involved in multiple events (such as bookings or arrests), and each event can include multiple charges, which makes it difficult to match information about one event. This is particularly true when one organization keeps information based on charges or events and another organizes its record system by people. One strategy that can be effective in dealing with multiple charges is to collect information about the most serious charge. If that charge is later dismissed by the court, then the data collector gathers information about the next most serious charge. This strategy works well for most planning processes.

To address this case-matching problem, an integrated criminal justice information system assigns person, event, and charge identifiers linked across agencies. This often means that a portion of the data collection may be manual. For example, the data collection may include analyzing case flow in the courts. The primary source of information is, however, the jail's record system. It may be more accurate to

develop a data collection sheet or format from the court records than to try to electronically match the records and manually enter the new information on the paper or electronic form.

Conclusion

Now that you have the information, the worst of the data collection is over. The next stages are exciting. Chapter 6 suggests some ways to analyze the **information** and introduces some statistics that are found in almost all statistical studies. Chapter 7 provides the opportunity to figure out what those statistics **really** mean.

Chapter 6: How to Analyze Information

Many people may develop an aversion to statistics, believing that:

- Only someone with a PhD can understand statistics,
- Understanding statistics requires expertise in calculus, and
- Statistics can be made to say whatever analysts want.

These popular misconceptions are reinforced by statistical jargon. No one can deny that at the level of proving and discovering theoretical relationships, the field of statistics is complex. But it is no more essential to understand the theoretical proofs than it is to understand why the internal combustion engine works to be able to drive! What is important is to apply the right statistical tools to the data being analyzed, understand some statistical assumptions, and be aware of the limitations of these techniques.

You should now put aside your assumptions and preconceptions and pretend you have never heard of statistics. This chapter presents a discussion of statistics that differs from most statistics courses. It will not spend much time explaining how to compute the statistics, and it will not require the memorization of any formulas. A computer can provide a wealth of statistics—so many, in fact, that some may have to be ignored.¹

This chapter:

- Defines statistics and identifies the purposes for which statistics can be used in jails,
- Explains the differences between statistics calculated for an entire population and statistics calculated for a sample (a specially selected subgroup) of that population,
- Defines the types of results of statistical computations,
- Introduces a number of common statistics, and
- Describes a number of statistical "sins."

What Is Statistics?

Statistics, as a field, is an area of mathematics that uses numbers to make large or diverse amounts of information more understandable. Most people can remember only limited amounts of information at one time. For example, when asked to summarize the findings about the entire inmate population of a jail for a year, most people cannot do so without using numbers to represent facts.

¹For those who need to compute more complex statistics than those explained in this chapter, a good resource for non-statisticians is a book cited in Appendix B, *How to Calculate Statistics*, by Carol Taylor Fitz-Gibbon and Lynn Lyons Moms. Another basic statistics textbook is *Understanding Statistics*, by William Mendenhall and Lyman On. Both of these sources have been used throughout this chapter.

The term "statistic" is also frequently used to mean the information that has been summarized by statistics (e.g., "jail statistics"). The potential for confusion is great enough without using the same word for two terms. For that reason, this manual uses the term "Data Element" for what many would call a jail statistic.

Basic statistics have three main purposes:

1. To summarize information,
2. To indicate how seriously to take differences between groups of people or events,
3. To help determine how strongly pieces of information are related to each other (Taylor Fitz-Gibbon & Morris, 1978).

Certain kinds of statistics are used for each of these purposes. This chapter deals with which statistics go with which purposes and which kind of data elements later. In the meantime, let's focus on one of the most important dividing lines in statistics—the difference between statistics based on entire populations and those that come from samples (parts of a whole population).

Statistically Significant Differences

In Chapter 2, this manual introduced an important term, "population," which, in statistics, means all the cases (persons, events, objects, times, etc.) about a data element that could be collected. In Chapter 5, another important term, "sample," was introduced. A sample is a specially selected part of a population. The difference between "population" and "sample" is so great that statisticians actually use different terms to refer to the statistics associated with populations ("parameters") and those that relate to samples ("statistics").

When data is collected about a population, and a parameter (such as a frequency distribution) is computed, analysts can be certain that the information is correct. For example, if an analyst collected Legal Status data for all the people booked at the jail during a given year and discovered that 80% of them were pre-trial and 20% were sentenced, those numbers are absolute.

However, if this information was collected from a sample of people booked at the jail, there is some uncertainty. A skeptic might ask, "How do you know that is true for everyone? You might have selected people for the sample who are somehow different from the 'real' jail population."

In statistics, when information from a sample is applied to the entire population, an "inference" is made. Analysts infer that what is discovered about a sample is equally true for the population. If the sample is collected randomly (by a sampling scheme like the ones identified in Chapter 5) and if it is large enough to adequately represent the population, making inferences is acceptable statistical practice, provided that the analyst identifies how reliable the inference is. The mechanism that allows analysts to make inferences about a whole population from a sample is probability.

Probability

Chapter 5 introduced the term "probability" as "the odds." An example will explain the concept more completely.

Let's revisit Sheriff #2. The Inmate Population Data at Sheriff #2's jail indicates that 60% of the people booked are pre-trial and 40% are sentenced. Sheriff #2 collected the Inmate Population Data over an entire year, when there were 2,500 bookings.

Sheriff #2 and the county commissioners are working together to build a new jail to replace the 60-year-old facility the county has outgrown. Inmate Profile Data is needed to help the architect with programming the facility, but Sheriff #2 is disturbed by the prospect of collecting Inmate Profile Data for 2,500 cases. A statistical solution allows Sheriff #2 to randomly select a sample of about 750 cases to represent the whole population, which is preferable to collecting data on 2,500 cases.

However, to be on the safe side, Sheriff #2 decides to experiment with the sampling procedures. He heard that some people questioned numbers reached by statistics in other counties building new jails. Because of the manual information system, Sheriff #2 knows what some of the actual population parameters are. Randomly selecting 100 of the last year's Inmate Information Cards, Sheriff #2 asks the jail administrator to count how many pre-trial and sentenced inmates are in the sample. Sheriff #2 repeats this process three times. The jail administrator tabulates the results, which are displayed in Table 6.1.

Sheriff #2 was convinced that sampling would work and that he could defend using it in the Inmate Profile Data Collection. Furthermore, he was convinced that once the probability of getting a specific sample was known, it would be relatively easy to decide whether to agree with the statistics. Sheriff #2 had grasped an important point about **probability**. Probability measures a belief in a particular outcome. Often, in statistics, probability is defined as the number of times a particular outcome will occur if a test or experiment is repeated many times. This is a criterion for deciding whether the results are credible. Probability theory is one of the foundations of statistics.

Table 6.1 Results of Sheriff #2's Experiment

Sample	Pre-Trial	Sentenced
Sample #1	79	21
Sample #2	77	12
Sample #3	82	18

Unfortunately, life is not always as easy as it was for Sheriff #2 in this example. Often, the true population parameter is unknown. Statisticians have provided tools that can be used to determine how sure we can be that statistics derived from a sample represent the whole population. However, they work only when the sample has been randomly selected. All other samples have unknown probabilities and are, as a result, useless for statistical purposes. Randomly selected samples have what is usually called "normal distribution" (or "bell-shaped distribution"), which is the other statistical foundation covered in this manual.

Normal (Bell-Shaped) Distribution

Normal or bell-shaped distribution is not just a statistical or mathematical phenomenon—it is real. Once it is understood, many aspects of statistics that were confusing may suddenly start to make sense. An example will help to illustrate this phenomenon. The jail administrator, who tabulated the results in the last example, was surprised by Sheriff #2's ability to pull 100 cards out of the information system that represented the whole inmate population so well.

To make sure Sheriff #2 had not used some kind of trick, the jail administrator tried an experiment. Thinking that Legal Status might be too easy to manipulate (after all, 60% of the inmates were pre-trial), the jail administrator chose a data element that would be much harder to "fix"—Length of Stay (LOS). The information system revealed that the average LOS was 20 days. Thinking that the sheriff's three attempts to test the sampling procedure were not enough, the jail administrator repeated the sampling procedure 100 times. Using the sheriff's procedure, the jail administrator randomly selected 100 cards from the information system, calculated the average Length of Stay for those 100 cards, and recorded the result on a piece of paper. The jail administrator then put those cards back and randomly selected another 100 cards. This continued 98 more times. When the results were posted, they looked like Figure 6.1.

Sixty-eight out of the 100 trials resulted in samples with an LOS between 18 and 22 days, and 95 out of the 100 trials resulted in samples between 16 and 24 days. By now, the jail administrator was a believer. What the jail administrator discovered is a phenomenon that characterizes nearly all data elements. If selected randomly and graphed like Figure 6.1, data elements are distributed to form a very distinctive curve called a "normal" or "bell-shaped curve."

What is most important in the jail administrator's experiment is the fact that randomly selected sample averages ("means") also have a normal distribution. This is a basic statistical assumption. It is what makes many statistics work. Statisticians and mathematicians discovered that this "normal" curve has some special mathematical properties, which are discussed later in this chapter.

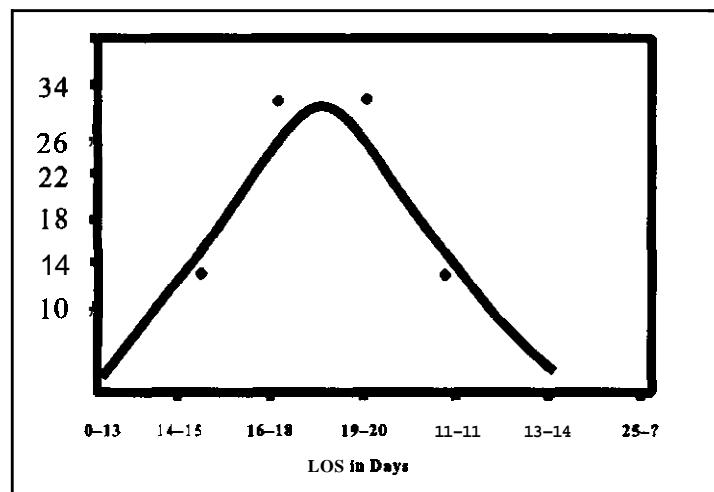
Right now, you deserve some credit, because in the last three and a half pages, you have absorbed two topics that often take half a semester in a basic statistics course. Undeniably, more detail is provided in basic statistics courses, but once you understand probability and normal distribution, you understand two major statistical concepts.

Statistical Results

Almost incidentally, the jail administrator's experiment illustrates two other important statistical facts. First, statistics can be expressed by using graphs and charts as well as numbers. In fact, reports in which statistics are used benefit by a good display of the data. Using graphs and charts effectively is so important that this manual devotes much of Chapter 8 to this topic.

The second concept the jail administrator discovered was discussed in some detail in Chapter 4: "Levels of Measurement." This is another basic concept that organizes the field of statistics. It refers to how different types of data elements can be measured. Data elements are divided into three basic categories. The first group (whose answers can be divided into a series of categories) is called nominal

Figure 6.1 The Jail Administrator's Experiment



data elements. The second group (whose answers can be ranked in some kind of order, like from good to bad, or high to low, etc.) is called ordinal data elements. The third group (whose answers are expressed in equal units) is called interval and ratio data elements. Statisticians divide the third group into two subgroups, depending on whether " 0 is a possible response. The level of measurement of each data element determines the types of statistics used to analyze it. These data element categories actually fall on a kind of scale themselves. This scale (or continuum) ranges from the simplest (nominal) to the most complex (interval/ratio).

The jail administrator discovered something it took statisticians some time to figure out: A more complex data element can be graphed or expressed like a simpler one. In the example, Length of Stay (which is an interval/ratio data element) was grouped by ranges of scores (cases with a LOS of 0–14 days, 14–16 days, 16–18 days, etc.). In doing so, interval/ratio data was turned into ordinal data. That, in turn, allowed the data to be displayed on a bar chart, which was easier to understand. It also would have allowed a statistician to use a different statistical procedure (a frequency distribution)—one that is much easier to use without a computer—to analyze the data.

Unfortunately, it is impossible to turn simpler data into more complex data. This is just one more compelling reason to collect data carefully.

Statisticians call the process of grouping data into various ranges "aggregating the data." The reverse is called "disaggregation." Experimenting with different ways to group data is one way analysis takes place. Generally, analysis begins with the data grouped in the largest general categories and proceeds by disaggregating it to show progressively finer levels of detail.

Types of Statistics

The rest of this chapter focuses mainly on the statistics that summarize information (descriptive statistics). Some time is also spent on the other statistical types (those that examine differences **between** groups or relationships between data elements). These probably will be included in reports prepared for you, and some may be encountered in Special Issue Data Collections analyzed on a computer.

Descriptive Statistics

There are two basic types of descriptive statistics:

1. Measures of central tendency (which locate the center of any kind of distribution of data) and
2. Measures of variability (which define how widely the data is spread).

Together, these two types of descriptive statistics can provide much information about the data with a minimum of math. With the exception of frequency distributions, which work with data elements of all levels of measurement, these statistics make sense only with interval and ratio level variables.

Type 1: Frequency distributions. Frequency distributions were discussed earlier in this manual, but they are worth a short review because they are among the most useful basic statistics. They are actually where all descriptive statistics begin. Frequency distributions are calculated by dividing all the data points (the scores or individual measurements for one data element or variable) into a series of categories (either

because they fall into categories naturally or because the data was divided into groups). The number of cases that fall into each group are counted and then expressed graphically (e.g., Figure 6.2) or in numbers (including both a percentage and the number of cases in each group). Table 6.2 is a frequency distribution displayed in percentages.

Figure 6.2 Frequency Distribution of Age

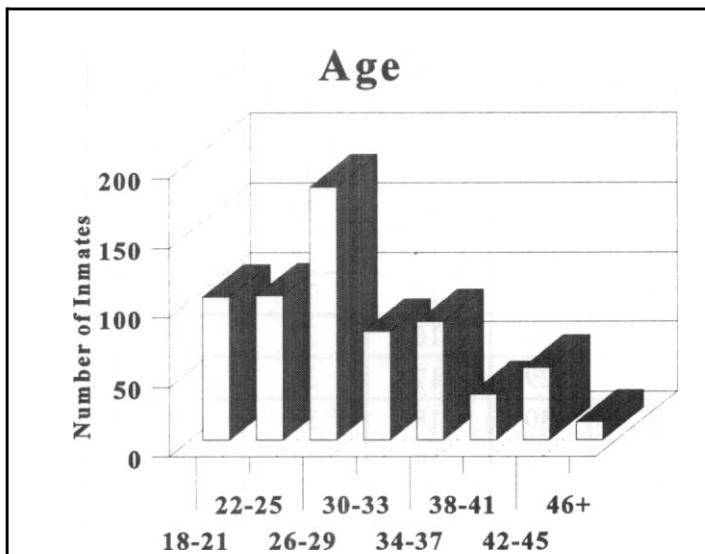


Table 6.2 Frequency Distribution of Age

Age	Number	Percentage
Under 18	0	0%
18-21	103	16%
22-25	104	16%
26-29	182	28%
30-33	78	12%
34-37	85	13%
38-41	33	5%
42-45	52	8%
46 or older	13	2%

An analyst can learn a lot about the jail population from just this frequency distribution. It is possible to tell:

- The percentages and numbers of people in various age groups in the jail,
- How many people were included in the sample, and
- That the population distribution leans toward the younger end of the age scale (statisticians would say that the distribution is skewed).

Sometimes, the data is neatly divided into categories, and the frequency distribution is obvious. Nominal and ordinal variables, like Legal Status or Charge Status, which have values that fall into distinct categories, are examples. However, **interval/ratio** variables have a wide range of possible values. To express these variables in a frequency distribution, the possible values must be divided into groups. Several **rules** can help.

Rule 1: Divide the possible values into groups that "make sense." Statisticians call these groups "classes." In the previous example, you should not divide the data about age into classes that consisted of 13 months.

Rule 2: Make the classes small enough so that the data is not bunched into one group. In the previous example, you should not divide the jail population into groups of 18 and younger, 19–50, and over 50.

Rule 3: The classes should generally be equal in size. For example, the same number of years should be included in each class.

Rule 4: No value should fall into two of the classes. For example, if you had two classes, 18–21 and 21–24, it would be impossible to know where to put someone who was exactly 21 years old. Pick boundaries for each of the classes so that no cases fall on the boundary line between two classes.

Type 2: Measures of central tendency. While frequency distributions are helpful, sometimes they are not very convenient and can take up a lot of room in a report. Descriptive statistics has another way to counteract these problems. Measures of central tendency locate the center of a frequency distribution.

1. **Central Tendency Measure 1: The Mean.** The mean is an arithmetic average. To calculate the mean, add all the measurements for a particular data element and divide the result by the total number of measurements in the set. If Sheriff #2 wanted to calculate the average daily population of the jail for a given month, he would add the counts for each day and divide by the number of days in the month (Table 6.3).

Table 6.3 Calculating the Mean Jail Population for One Month

Day Number	Count	Day Number	Count	Day Number	Count
1	26	11	36	21	34
2	32	12	31	22	29
3	37	13	32	23	27
4	33	14	29	24	28
5	29	15	25	25	31
6	28	16	28	26	33
7	29	17	30	27	30
8	30	18	33	28	35
9	37	19	34	29	31
10	35	20	31	30	29
Total = 932					
Mean = 932 / 30 = 31.07					

When most people talk about an "average," this is what they mean. However, two other measures of central tendency can be used in addition to (or instead of) the mean. (That "instead of" will receive a little extra attention in the section of this chapter on statistical sins.) Sheriff #2 did not stop with just an average, but continued to explore several other measures of central tendency.

2. **Central Tendency Measure 2: The Median.** The median is the middle measurement when all of the measurements in the set are arranged according to size. Statisticians call this process "rank ordering." If there is an odd number of measurements in the set, the median falls exactly on the middle measurement. If there is an even number in the set, the median falls exactly halfway between the middle two measurements. Using the same data, Sheriff #2 put all the measurements in order from lowest to highest. Table 6.4 is the result.

Sheriff #2 noticed that although the median (31) and the mean (31.07) were not very different, neither were they exactly the same. He thought that the difference was not caused by a math error, because the two measures of **central** tendency actually measure different things, so he calculated one more measure of central tendency.

Table 6.4 Calculating the Median

Day Number	Count	Rank Order
3	37	1
9	37	2
11	36	3
10	35	4
28	35	5
21	34	6
19	34	7
26	33	8
4	33	9
18	33	10
2	32	11
13	32	12
29	31	13
12	31	14
20	31	15
25	31	16
17	30	17
27	30	18
8	30	19
30	29	20
14	29	21
22	29	22
7	29	23
5	29	24
6	28	25
24	28	26
16	28	27
23	27	28
1	26	29
15	25	30

➡ = The Median

3. **Central Tendency Measure 3: The Mode.** The mode is the measurement that occurs most frequently in a set of measurements. To calculate the mode, Sheriff #2 arranged the scores like a frequency distribution. Table 6.5 shows the results.

As might be expected, the mode (**29**) is different from both the median (31) and the mean (31.07). In this case, all three measures of central tendency are very close, which would lead an analyst to believe that the data is both normally distributed and clustered closely together. However, this is not always the case. Sheriff #1's jail was about the same size, but the jail population had greater fluctuations. And in a nearby county, Sheriff #3 had a small jail with some unusual population problems. Table 6.6 compares the Daily Counts for the three jails during the same month.

Table 6.5 Calculating the Mode

Count	Number of Occurrences
25	1
26	1
27	1
28	3
29	5
30	3
31	4
32	2
33	3
34	2
35	2
36	1
37	2

➡ = The Mode

Table 6.6 Measures of Central Tendency at Three Small Jails

Day No.	Jail #1	Jail #2	Jail #3
1	47	26	14
2	44	32	15
3	30	37	16
4	28	33	18
5	15	29	17
6	21	28	16
7	27	29	13
8	49	30	19
9	50	37	15
10	24	35	17
11	19	36	16
12	10	31	13
13	12	32	19
14	24	29	15
15	38	25	17
16	40	28	16
17	46	30	15
18	36	33	17
19	33	34	18
20	31	31	14
21	26	34	16
22	43	29	172
23	40	27	167
24	23	28	161
25	28	31	16
26	25	33	18
27	21	30	14
28	20	35	17
29	43	31	15
30	39	29	16
Resident Days	932	932	932
Mean	31.07	31.07	31.07
Median	29	31	16
Mode	44	29	16

The table clearly shows why it is important to look at more than one measure of central tendency. Extreme cases can distort the mean, but the median and the mode are less likely to be affected by extreme cases. As you can see, the three jails have very different patterns of use.

Sheriff #1 has a problem with weekenders. The numbers in Jail #1's column regularly peak every seven days.

- Sheriff #2 has an evenly distributed population. It appears that the judge who sentences most people supports the sheriffs work program and does not use weekend sentences.
- Sheriff #3's county is generally idyllic, rural, and heavily forested, with a low population. There were few major law enforcement problems until a motorcycle gang camped there during a long holiday weekend. Sheriff #3 had a hectic weekend, and the jail statistics for the month reflect it.

This is an example of a major problem with the measures of central tendency. Without another statistic that measures how much variation exists in the data, all of the measures of central tendency are incomplete.

Type 3: Measures of variability. Measures of variability tell how widely the data is distributed. If the measures of central tendency tell where the middle of the bell-shaped curve is, the measures of variability describe how **wide** the curve is from side to side. There are four important statistical measures of variability.

1. **Measure of Variability 1: Range.** Range is the simplest measure of variability to calculate and use. In a series of data points (individual scores or values), the range is the difference between the highest and lowest data points. In the case of the three jails, the range is quite different for all three. Table 6.7 illustrates the difference.

Table 6.7 Range at the Three Jails

Jail No.	High Count	Low Count	Range
1	50	10	40
2	37	25	12
3	172	13	159

Range is also the fastest measure of variability to calculate. By looking at the high score, the low score, and a measure of central tendency, it is possible to tell if all the scores are grouped close together or if the scores are widely scattered. The more the scores are scattered, the less confidence should be placed in the average.

2. **Measure of Variability 2: Percentiles or Quartiles.** Percentiles refine the median (the middle measurement when all of the measurements in the set are arranged according to size). They are often used to divide all the scores into 10 groups (i.e., so that scores are located in the 80th, 90th, etc., percentile). Quartiles divide all the scores into four equal groups. The quarter that is the lowest is below the 25th percentile; scores that are below the median are below the 50th percentile, etc. The score exactly between the median and the highest score becomes the upper quartile (75% of all the scores are below it). The score exactly between the lowest score and the median (25% of all scores are below it) becomes the lower quartile.

Percentiles and quartiles provide a good view of how the data is distributed. They are frequently used to report scores for tests such as the Scholastic Aptitude Test. Although they are seldom used in jail data, they are useful in beginning to analyze the data. A disadvantage is that quartiles and percentiles do not define the highest and lowest scores.

3. **Measure of Variability 3: Variance.** Variance measures how far individual data points lie from the mean. It is rarely mentioned in anything but technical papers, but is essential for calculating probably the most famous measure of variation, the standard deviation.
4. **Measure of Variability 4: Standard Deviation.** The standard deviation is the square root of the variance. Several formulas for calculating the standard deviation (including a shortcut) are included in Appendix I.* The most important thing to understand is how to use and interpret a standard deviation.

What is important about the standard deviation? If distribution is normal (the data is distributed in a normal or bell-shaped curve) and one standard deviation is added to and subtracted from the mean, 68% of all the cases will be found in the intervening interval. If approximately two standard deviations — 1.96 standard deviations to be exact—are added to and subtracted from the mean, 95% of all cases will be found in the interval—this represents a 95% confidence interval. If three standard deviations are added to and subtracted from the mean, then nearly all (99%) of the cases will be contained in that interval. Statisticians call this the "Empirical Rule." This phenomenon provides a great deal of information to an analyst from just two statistics, the mean and the standard deviation. Not only can the central point in the data be identified, but an interval can be placed around it that defines where 95% of all the cases in the data set are found.

Before closing this section on descriptive statistics, let's calculate the standard deviation for the Daily Counts of the three jails we have been studying (Table 6.8). Remember—the mean for all three jails is 31.07.

Table 6.8 Standard Deviations of **the** Three Jails

Jail No.	High Count	Low Count	Range	Std Dev
1	50	10	40	11.12
2	37	25	12	3.10
3	172	13	159	45.25

With the standard deviation calculated, the differences between the three jails are obvious when the Empirical Rule is applied. The Empirical Rule is that 68% of all cases lie within plus and minus 1 standard deviation, and 95% of all cases lie within plus and minus approximately 2 standard deviations. Sheriff #2's jail has the smallest variation in its inmate population. Sheriff #1's jail has a much broader range of daily populations, and, for that size, the standard deviation is large. Sheriff #3 has what seems to be the strangest situation of all. The standard deviation for the jail is larger than its Average Daily Population. However, this is relatively common in some kinds of jail data--especially Length of Stay data. To the experienced analyst, Sheriff #3's strange standard deviation means one of two things:

1. Two groups are represented in the data (one near the low end of the range and one near the high end) and few, if any, cases are in the middle. Statisticians call this bimodal distribution, which means that there are **two** normal curves (and two populations) instead of one.

*This will be useful if you do not have access to a computer. With a computer, you can use the statistical functions of your software to do the calculation.

2. Several extreme cases throw off the calculations. On some days (when the motorcycle gang was in town), Sheriff #3 had over 150 inmates in the jail.

The standard deviation is a great help to statisticians because it is one way to determine whether the distribution of the data is normal (and that is the basis for a great many statistical assumptions). The standard deviation is a kind of check on many basic assumptions.

This is about as complicated as basic jail statistics get, and if sheriffs and jail administrators use these tools to make good management decisions and to monitor the jail population, this manual will have achieved one of its primary goals. However, reports, articles, newspaper stories, etc., often cite other kinds of statistics. It is important to understand something about more advanced statistics and to know how to interpret them.

This manual also provides some information below about statistics that serve other purposes, but does not give much detail or discuss how to calculate them. A resource listed in Appendix B, *How to Calculate Statistics*, will do precisely that.

Statistics for Examining Differences Between Groups

Statistics that examine the differences between two or more groups are called Tests of Significance. Sheriffs and jail administrators are most likely to encounter significance tests in journal or magazine articles about inmate programs and services that claim to make a difference in inmate behavior either before or after release. Significance tests typically are used to determine differences in the number of arrests after release between two groups, one that participated in a pre-release program and one that did not.

What is significant about statistical significance? Many tests of significance work by using two of the basic statistical concepts, normal distribution and the Empirical Rule. Statisticians apply the Empirical Rule in a situation in which interval/ratio-level scores of two groups are compared. One group's score is allowed to represent the mean. Using logic, statisticians deduce that if the other group's score falls within the intervals defined by the Empirical Rule, the difference might be attributed to the sample, and therefore the difference might not be "real." Therefore, these differences are not "statistically significant."

The opposite is also true. Using one group's score as the mean, if the other group's score falls outside the interval defined by the Empirical Rule, the difference might be "real" and not merely a sampling difference. If the other score is outside the two standard deviations in either direction from the mean that contains 95% of all the cases, it has to be in the other 5 percent of the cases. This kind of difference is statistically significant at the .05 level. For those who are not satisfied with significance at the .05 level, mathematicians determined that to be significant at the .01 level, the interval around the mean had to be plus and minus 2.58 standard deviations. That increases the odds to 1 out of 100 times that the difference might be real and not just a statistical quirk.

Errors in statistics. Notice that the previous discussion is full of "mights." Statisticians recognize the possibility that the difference might be a "statistical long-shot" and their logic might be wrong. Statistical logic can be wrong two ways, called Type I and Type II Errors.

Type I Error occurs when researchers say there has been a change when there has not been. The difference that was detected is just one of those long-shot sampling flukes. The sample that was drawn was one out of the 5% or even 1% of all the possible samples that is different from the real population. Type I Error can be controlled by using small significance levels. Type II Error occurs when researchers say there has not been a change when there really has been. Program evaluation results might suggest that the difference **between** two groups is significant only at the .15 level. Traditional researchers would say that this difference is "not statistically significant" and might discontinue the project.

Generally, the best prevention for Type I and Type II Errors is good sampling procedures and large enough samples. But this problem has led to a major debate in social research. Since our understanding of what makes people change is imperfect, it is possible that the measurements used to evaluate programs that are supposed to change people are not able to pick up true differences. Researchers might say that a program has not made a difference simply because they cannot measure accurately what has happened.

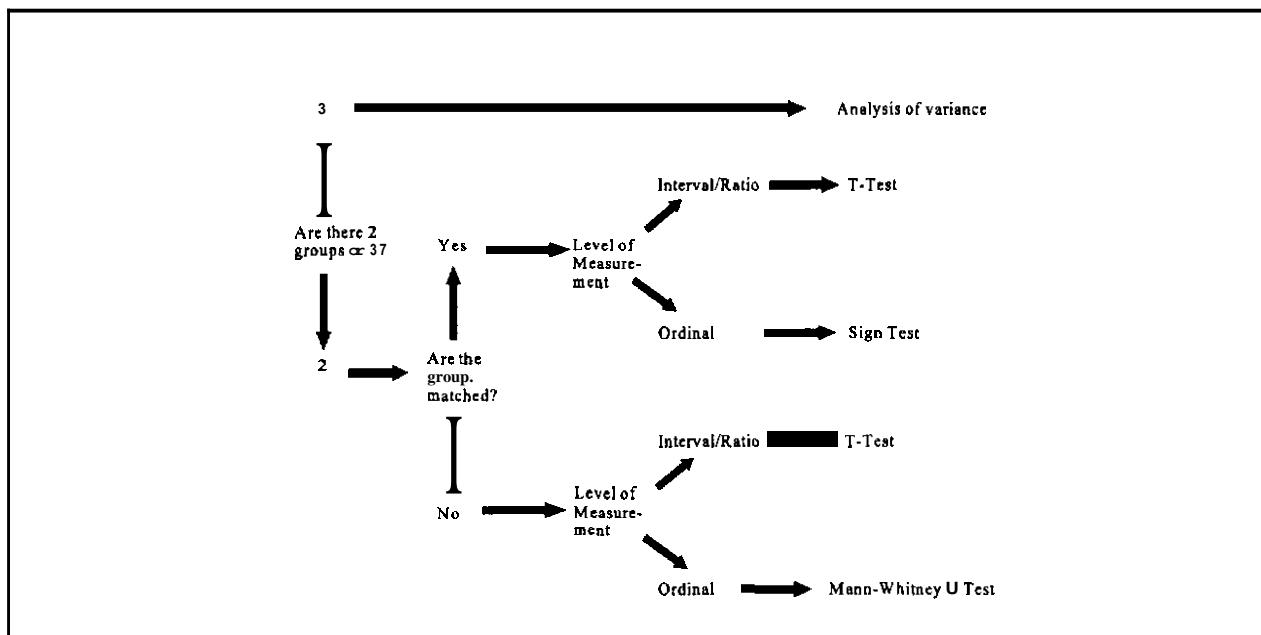
Analysts must use some judgment in this area. Significance levels are only tools that sometimes may be helpful in making management decisions. For example, consider this situation. One group of inmates who participated in a pre-release program at the jail had an average of **4.3** arrests after release; another group who did not participate had **4.4** arrests after release. If the groups were very large, these results could be statistically significant at the .001 level. But the real difference between the scores is so small that the results are not significant from a management perspective. However, if the situation were changed so that the differences were significant only at the .15 level, but the difference in arrests were much greater (e.g., 1.3 arrests for the group that participated in the pre-release program compared with **4.4** for the group that did not), sheriffs and jail administrators are likely to have a better opinion of the pre-release program.

Deciding which kind of significance test. There are many different tests of statistical significance. Variations can be used with different numbers of groups, different sizes of groups, different types of groups, and different types of data. Other tests include:

- T-Tests,
- Mann-Whitney U-Tests,
- The Sign Test, and
- Analysis of Variance.

A technique that researchers call a "decision tree" helps you to decide which test to use. This involves working through a series of questions that ultimately lead to the only possible answer. Figure **6.3** illustrates a decision tree to determine which significance test should be used.¹

¹Figure 6.3 is based on a decision tree developed in How to Calculate Statistics, which also provides the instructions for actually doing the significance tests mentioned here and develops worksheets that will help in their computation.

Figure 6.3 Test of Significance Decision Tree

Question 1: How many groups are there? If three or more groups are included, an advanced statistical technique called Analysis of Variance will be needed to measure the differences in scores among groups. If analysis of variance is required, expert help probably will be necessary as it is a complex statistical procedure. If only two groups are included, other statistics mentioned here can be used.

Question 2: Are the groups "matched"? Researchers consider groups "matched" if each case in Group I is paired with a similar case in Group II. A "match" can be a person with similar characteristics, or it may be the same person who has been tested at another time. A likely instance of "matched groups" in correctional research is when one group is compared before and after participation in a program.

All other groups have to be considered "unmatched." When working with unmatched groups, the relative sizes of the two groups are important, as is the amount of variation within each group. These two factors will help to determine how different their scores are likely to be. The smaller the groups or the more variation in the scores, the more likely that one group's mean will be different from the mean for both groups together.

Question 3: What level of measurement are the scores? Significance tests that are used on the wrong kind of data are worthless. Differences between nominal level variables (data elements that are divided into discrete categories, e.g., the Pre-Trial and Sentenced categories of Legal Status) show the relationship between the two categories that are being measured. This chapter discusses this kind of relationship later. Differences between ordinal-level variables (data elements that are rank-ordered, e.g., arranged from lowest to highest) are measured by two tests. If the groups are not matched, the Mann-Whitney U-Test should be used; if they are matched, the Sign Test can be used. Differences in interval/ratio-level variables (data elements like age, number of miles, dollar costs, etc., that have equal units on a scale) are measured by T-Tests. T-Tests have been developed for both matched or unmatched groups.

Most of the statistical tests of difference are not difficult to calculate, but are merely tedious and involve a great many repetitions of similar mathematical functions. If you must calculate these tests by hand, a good programmable calculator (or, better yet, a statistical calculator that allows you to enter the data and then does the test for you) would be a great asset. Most special data collections are automated, and the computer can print out the results of these tests with the significance levels that are attached to them. With the understanding of statistical significance gained in this chapter, you should be ready to handle most of the common statistics of this type that you encounter.

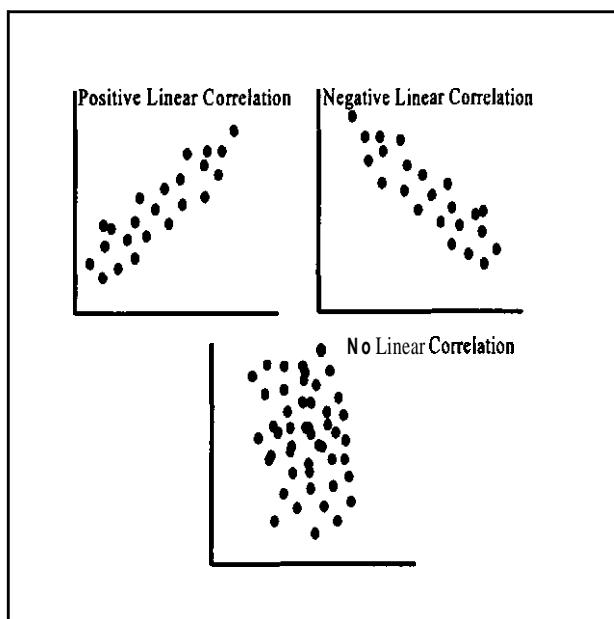
Statistics to Examine the Relationship Between Data Elements

Many statistics test the relationships, or "correlations," between data elements. Correlations are a statistical shorthand that describe the nature and strength of relationships **between** data elements.

Types of relationships between data elements. Correlations operate **in** a variety of ways:

1. The more one variable increases, the more another variable increases (e.g., the greater the number of miles driven on a transport, the greater the number of gallons of gasoline used);
2. The more one variable decreases, the more another variable decreases (e.g., the fewer the number of prior arrests, the lower the level at which bond is set); and
3. The more one variable increases, the more another variable decreases (e.g., the older the prisoner, the fewer the **number** of incidents in which he/she is involved at the jail).

Figure 6.4 Data Clusters



Correlations like those described in items 1 and 2 above are called positive correlations because a change in one direction in one variable is accompanied by a change in the same direction in another variable. The variables have a direct relationship; statisticians abbreviate correlations like these by the symbol "+r." Correlations like the one described in item 3 are called negative correlations because a change in one direction in one variable is accompanied by a change in the other direction in another variable. The variables have an inverse relationship; statisticians abbreviate correlations like these by "-r."

All correlations are expressed in numbers between +1 and -1. The "+" sign denotes positive correlations while the "-" sign denotes negative correlations. Numbers that are close to either +1 or -1 (i.e., $\pm .7$ and above) indicate very strong relationships.

If a graph showing all the data points for a strong positive or negative correlation is plotted, the data points "cluster" together. A straight line can be drawn through the cluster that "summarizes" the cluster. When it is relatively easy to draw this straight line, the relationship between the data elements is said to be "linear." This is called the trend line, which is used by statisticians for a special kind of correlation

called population forecasting or trend analysis.¹ In population forecasting, statisticians attempt to determine if there is a relationship between a data element, such as Average Daily Population, and time. Figure 6.4 provides examples of data that show different types of correlation.

Actually seeing a plot of the data points is important because there are times when correlations should not be used to measure the relationship between variables. The correlation itself cannot reveal this problem. In fact, some very strong relationships that are not linear exhibit weak correlations. For example, when taking tests, people who have either very low or very high anxiety levels tend to get low scores, but people who have moderate levels of anxiety get high scores. Relationships like this are called curvilinear. While there are few in basic jail data, they are possible. The best way to detect them is to graph the data. If such a relationship does emerge from the data, use a statistic called a Chi Square or Phi to find out if there is a relationship between the data elements.

Statistical tests for correlation. Most statistical tests for correlation are mathematically descended from a test called the Pearson's Product Moment Correlation. Common statistical tests of correlation include:

- Phi Coefficient,
- Rank Biserial Correlation Coefficient,
- Point Biserial Correlation Coefficient,
- Spearman's Rank Order Correlation Coefficient, and
- Pearson's Product Moment Correlation Coefficient.

Another statistic used to assess a relationship that is not a **true** correlation is the Chi Square. A series of questions will help to decide which test to use (Taylor Fitz-Gibbon and Morris, 1978).

Question 1: What kind of measurement is used for the first data element measured? Is the case classified (put into categories), ranked (put in some order), or rated (given a score)? There are no problems if the measurement is ordinal or interval. However, if the case is classified (nominal), the researcher must answer question 1a before going to question 2.

Question 1a: Does the measurement of the data element put the case into two categories or more than two? If the answer is more than two (as it will be for many jail data elements), a correlation cannot be used to define the relationship between the variables. However, a Chi Square can be used to test for a relationship.

Question 2: What kind of measurement is used for the second data element measured? See question 1. If the case is classified, the researcher also must ask Question 1a for the second data element.

Table 6.9, reprinted from Taylor Fitz-Gibbon and Morris, shows which test of correlation should be used, depending on the results of the answers to these questions.

¹This advanced statistical technique is beyond the scope of this manual. Most jurisdictions require technical help from someone experienced in this method.

Table 6.9 Appropriate Correlation Coefficients for Use with Different Types of Measurements

Classification of First Measure	Classification of Second Measure		
	Dichotomous	Ordinal	Interval
Dichotomous Examples: Yes—No Agree—Disagree U.S. Citizen—Alien Sentenced—Not Sentenced	Phi Coefficient	Rank Biserial	Point Biserial
Ordinal Examples: Rank in Class Low, Moderate, High Never, Sometimes, Often Negative, Neutral, Positive	Rank Biserial	Spearman's Rank Order	Spearman's Rank Order
Interval Examples: Arithmetic Score Length of Stay Years of Education	Point Biserial	Spearman's Rank Order	Pearson's Product Moment

Another statistical relationship test: The Chi Square. Correlations cannot be used when either of the variables is classified and put into more than two groups. The statistical test that allows analysts to identify whether being a member of a particular category of one data element has anything to do with being a member of a particular category of another data element is the Chi Square.

This is a useful statistical test for sheriffs and jail administrators because most of the basic jail data elements can be categorized and their values usually fall into more than one category. It is also easy to calculate. The Chi Square begins with what statisticians call a "contingency table" or a cross-tabulation, which is a chart that displays the relationship between two frequency distributions. The Chi Square test indicates whether the results of the contingency table are different from the results that would be expected if the two data elements were not related. Appendix J provides step-by-step instructions for calculating the Chi Square.

Interpreting measures of relationship. It is important to be cautious when interpreting correlations and Chi Squares. One common problem in interpreting these statistics is the fact that people assume that one data element has caused the other because a relationship exists between them. Correlation is not the same thing as causation. Think about height and weight, which have a strong positive correlation. But does height cause weight?

The technical aspects of this section have been included as resources, but are not part of the basic descriptive statistics that are tools for better management. If sheriffs and jail administrators can begin to keep good descriptive statistics on the jail population, then a major goal of this manual will have been achieved. The discussions about the statistics themselves should make sheriffs and jail administrators much more informed consumers of the statistics they receive.

Statistical Sins

Statistical sins can be sins of either commission or omission, and both can leave the statistical consumer in a very uncomfortable position. Part of the problem is that sometimes the most carefully calculated statistics are used carelessly; they may be interpreted by people who really do not understand statistical foundations. An additional purpose for this manual is to make sheriffs and jail administrators more knowledgeable consumers of the statistics that they encounter on a daily basis, both on and off the job.

One of the best resources in this area is the book *How to Lie with Statistics*, written by Darrell Huff in 1954. Although the examples this book used to illustrate the "statistical sins" in it are dated, the sins Huff identifies are just as prevalent today as they were then. This book is the source from which the following section was drawn.

Sin 1: How could the results be wrong if the calculations were perfect?

Chapter 5 emphasized random sampling and this chapter showed why it is important. But, sometimes, in spite of the best intentions, samples can acquire a "built-in bias." This can happen in a number of ways:

1. Researchers may report as fact what respondents say they do. People have a tendency to either embellish or "play down" some aspects of their behavior. For example, they may report earning more money than they actually do.
2. Researchers may fail to mention how many people did not respond to their survey. Frequently, people with certain characteristics (perhaps some of those being measured) will fail to respond to a survey. College surveys that report how successful their alumni are may be guilty of this kind of built-in bias. Possibly alumni who are not very successful did not respond.
3. Researchers may ask questions about behavior that is not socially acceptable. Surveys that ask "Do you cheat on your taxes?" will probably discover that no one cheats on their taxes.
4. The source on which the sample is based does not represent the whole population. A frequent source of names for many surveys is the telephone book or a list of registered voters. While more people have telephones now than in the era when social research was beginning, some people still do not have telephones. This biases the sample. Even though the people who were surveyed were selected randomly, they were selected from a biased source.
5. If information is collected by interview, the time of day and location of the interviews can bias the sample. For instance, house-to-house, daytime interviews can bias the sample by filling it with people who do not work outside the home, and "man in the street" interviews miss the people who stay home.
6. The person asking the question can bias the response. Gender or verbal intonation of the interviewer may elicit responses that the respondent believes the interviewer is seeking.

In developing a sample, not only must the cases be randomly selected, but the sample must represent the whole population. This can be a problem in jail surveys. For example, a systematic random sample of all people booked at the jail was developed for an Inmate Profile Data Collection. A major concern was the number of prisoners who were management problems during booking (e.g., they were violent, incapacitated by alcohol, mentally ill, etc.). When the statistics were analyzed, the results seemed low to all who worked in the system. The problem was soon discovered. Some prisoners who were the greatest

management problems in the Booking Room were never booked. They were held on mental health detainees and not charged with a crime. No crime...no booking; it resulted in a biased sample.

Sin 2: What is the real average?

In the discussion of measures of central tendency, **three** different measures (mean, median, and mode) were presented. These are not always the same thing, as illustrated by Table 6.6. It is fine that these measures are not the same because they measure different things. The problem comes when researchers do not say **which** average they are using. Data elements that are expressed in dollars, like bond amounts, are most vulnerable to this kind of statistical misrepresentation. Figure 6.5 illustrates why.

Imagine how these differences could be used during a political campaign with a "get tough on crime" political platform. The incumbent says, "The average amount at which I set bond was \$750." The challenger says, "The incumbent is soft on crime. Why, the average amount at which bond was set was only \$200!" The different averages present different pictures of bonding practices. Watch out for "averages" that do not say which kind of average they are.

Sin 3: What is missing from this picture?

This statistical sin is illustrated by the data from the three jails. There are two aspects of descriptive statistics: measures of central tendency and measures of variability. The two of them should be inseparable. Frequently, however, the "average" is presented without any measure of variability, such as range or standard deviation. Recall Sheriff #3's weekend when the jail (with a usual Average Daily Population [ADP] of 16) was disrupted by 172 bikers, resulting in an ADP of 31, twice what the ADP was the other 27 days in that month. Mistaken conclusions could easily be drawn by using an "average" without something that tells how representative that number is. That is what the measure of variability does. Be a little skeptical of averages that stop before giving a measure of variability.

Another statistic that is frequently missing—the number of cases—is somewhat less obvious. And it may be more devious because it misrepresents one of the basics, the frequency distribution. Sometimes the data is presented as a percentage without the number of cases in each group.

Compare Table 6.10, which shows the number of cases in the whole sample (1,000) and how many cases were in each group, with Table 6.11, which gives only the percentage. In Table 6.11, there is no way to know if there are 100 or 1,000 inmates in the sample. Just remember that "50% of the cases" can be just one case, and be suspicious of results reported only in percentages.

Figure 6.5 Measures of Central Tendency

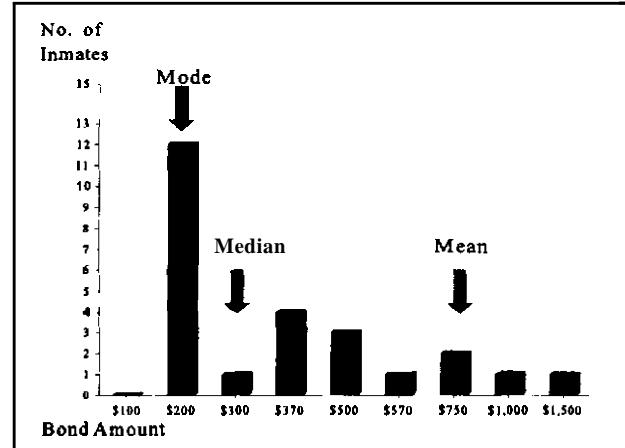


Table 6.10 Frequency Distribution of Age, #1

Age	Number	Percentage
Under 18	0	0%
18-21	160	16%
22-25	160	16%
26-29	280	28%
30-33	120	12%
34-37	130	13%
38-41	50	5%
42-45	80	8%
46 or older	20	2%
	1,000	100%

Table 6.11 Frequency Distribution of Age, #2

Age	Percentage
Under 18	0%
18-21	16%
22-25	16%
26-29	28%
30-33	12%
34-37	13%
38-41	5%
42-45	8%
46 or older	2%
	100%

Sin 4: So what if it is statistically significant?

With a very large sample, very small differences between groups can be statistically significant simply because there are many cases. From a manager's perspective, what difference does a pre-release program make if program participants average 4.21 subsequent arrests versus nonparticipants who average 4.23 subsequent arrests? This distinction may be statistically significant, but it would be difficult to use as justification for program funding.

Informed consumers look not only at whether the difference is statistically significant but also at what the real difference is. It is important to remember that some differences that are not statistically significant may be important for management decisions.

Sin 5: If A. then B?

This is so important that, although it was mentioned earlier, it is worth repeating: Correlation is not the same as causation. Just because one thing happens **before** another does, it does not imply that one causes the other. Strong correlations can result from:

1. Both factors being related to a third (and possibly unknown factor) that causes **both** of them.
2. Chance or "spurious correlations." An example of this is finding a strong positive correlation between a county's jail bookings and the price of hog futures on a local commodity exchange.
3. An actual relationship exists between two data elements, but it is not known which data element causes the other. One example would be a relationship between Length of Stay and the number of incidents committed by prisoners. Do prisoners commit more incidents because they stay longer and thus have more opportunity? **Or** does the number of incidents committed by prisoners predispose the court to keep bond high in the case of pre-trial inmates and reduce the amount of good time earned or increase the length of the sentence in the case of sentenced prisoners? **Or** are

long-term prisoners more likely to commit incidents than prisoners who are released within shorter periods?

Sin 6: What is wrong with this picture?

Statistics can result in both a numerical summary of information and a graphic representation. However, some of the worst statistical misrepresentations occur in how information is displayed. By carefully choosing the scale on which **information** is graphed, a strong (and sometimes misleading) message of what the data means can be depicted. Figures 6.6, 6.7, and 6.8 display the same data on three scales.

Figure 6.6 Scale Version #1

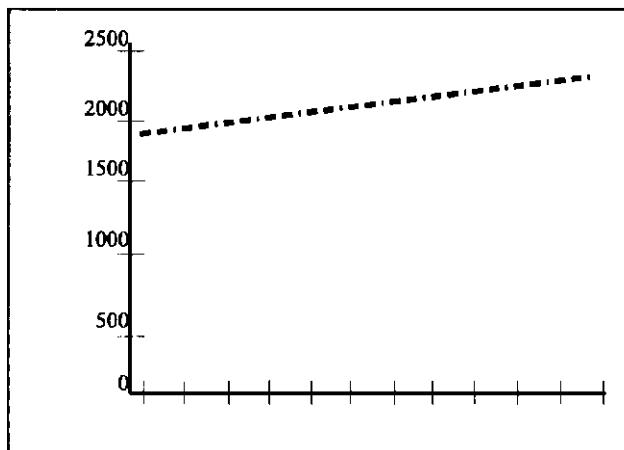


Figure 6.7 Scale Version #2

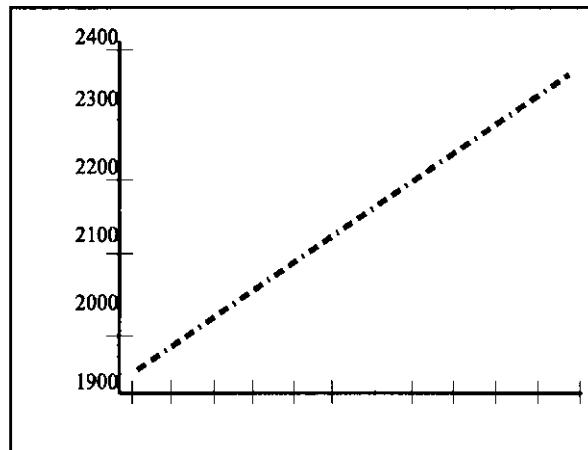
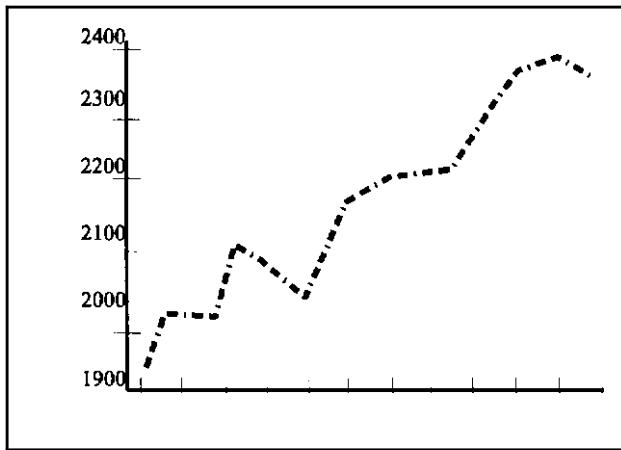


Figure 6.8 Scale Version #3



Most people react differently to the three graphs but do not look closely enough at graphs and charts to see the "tricks" that were played. Look at the numbers at the left of Figures 6.6 and 6.7. The scale in Figure 6.6 begins at zero; the scale in Figure 6.7 begins at 1900. Now compare the lines in Figures 6.6

and 6.7 with the line in Figure 6.8. Notice how many points are included in 6.7 and 6.8. Displaying data effectively and responsibly is a major part of any data collection. Sometimes there is a fine line between effective and ethical data display.

Conclusion

This chapter covered a great deal—from some of the foundations on which the entire field rests, to how to use the different types of statistics (and how to calculate some of the more basic ones), and finally to the statistical sins that have given the field a bad reputation. This manual should encourage sheriffs and jail administrators to keep jail statistics and use them to make more informed decisions about situations that affect the jail. But, learning how to understand and calculate statistics is only a part of this manual's mission. If sheriffs and jail **administrators** are to benefit from this knowledge, they need to learn how to apply it in their own situations and be able to interpret statistics. That is the purpose of Chapter 7.

Chapter 7: How to Interpret Information

When the information has been collected and processed, the exciting part of the data collection—the analysis itself—begins. Analysis is the process of interpreting what the information **means** in the context of the environment in which it was collected. Explaining how to begin this process is like telling someone how to ride a bicycle. It is possible to explain about balance, velocity, and the laws of motion, but until he/she actually starts pedaling, none of it makes much sense.

Ideally, sheriffs and jail administrators who want to learn how to interpret statistical information should work with someone who can help to analyze the data. The next best alternative is to use a "programmed text" that teaches analytical skills. In a programmed text, the reader is presented with a problem, given an opportunity to respond to it, and then given instructions to see the preferred response. Although you may be tempted to gloss over them, take the time to write down your responses. Writing helps many people think things through.

This chapter is structured as a programmed text. NIC can provide this part of the text on disk. Using WordPerfect 6.1, you can work through the examples that are provided in this chapter. Two case studies provide real-life examples of how basic jail data can be used for management decisionmaking. What the key actors chose to do in each case may not be exactly what **you** would do in the same position. Although you may not agree with their choices, these case studies make it clear how people in these **actual** situations used the information that was available to make critical policy decisions for their jails.

Case Study One: The Friday Night Blues

The Situation

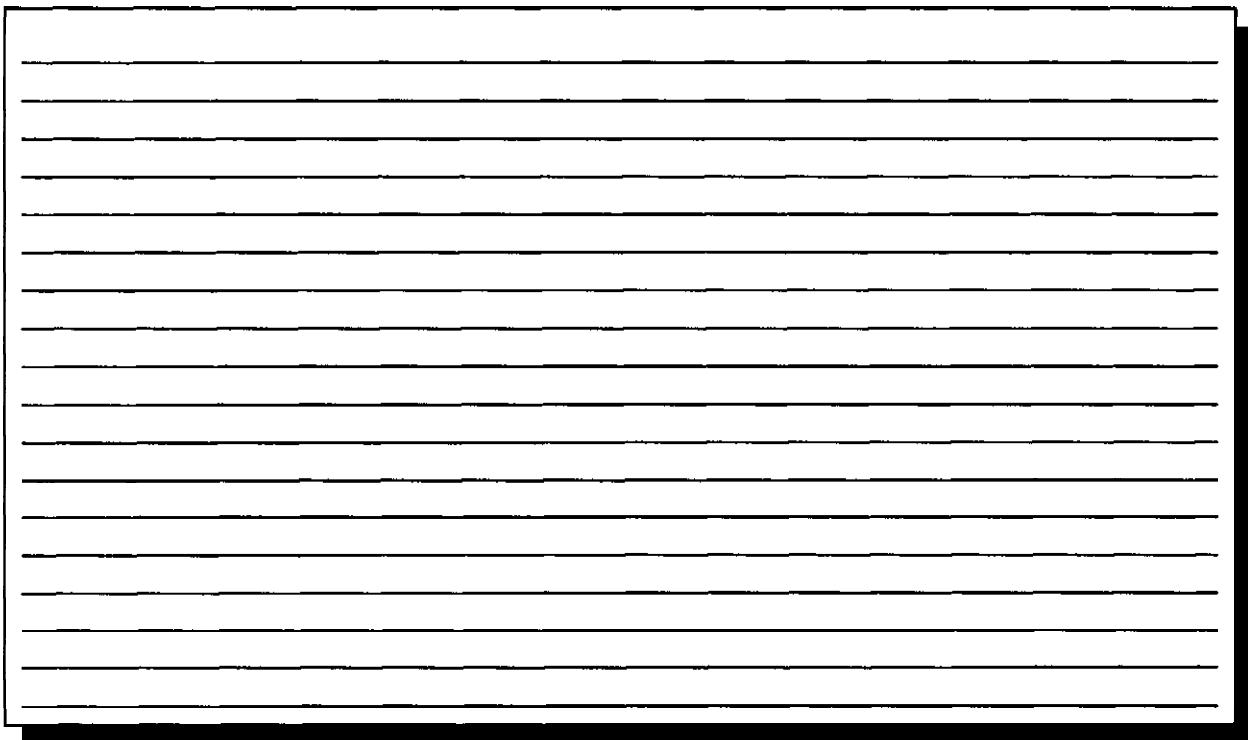
The Metro County Jail holds people who are intoxicated but not charged with an offense under two circumstances:

1. If they are violent or
2. If the local detoxification center is full.

When they are brought to the jail, they are held in the Booking Room for 8 hours and then released. Although the detox center has been operating at only about 50% of capacity, the number of detox holds at Metro County Jail has steadily increased over the 3 years that this practice has been followed. The booking area is frequently crowded and congested. If the situation gets worse, the booking area may need to be expanded.

You are the jail administrator. A special task force has been established by the sheriff and county commissioners to decide whether Metro County should expand the present booking area. The sheriff told you to have data available for the next meeting of the task force.

What should the jail administrator do first? (Write your answer below.)



A large rectangular box with a black border, designed to look like a sheet of lined paper. Inside the box are 20 evenly spaced horizontal lines for handwriting practice.

When you have completed your answer, go to the next page to see what happened.

The jail administrator and the sheriffs administrative assistant developed a problem statement below that described the situation. If you had a different response, you might want to review Chapter 2.

The Metro County Jail has the authority to detain intoxicated people without charges in two circumstances:

1. If the person is violent or
 2. If the detox center is full.

We hold these people in the Booking Room for 8 hours and then release them. The detox center has been operating consistently at about 50% of capacity, but the number of detox holds at Metro County Jail has steadily increased over the 3 years that this practice has been followed. The booking area is frequently crowded. If the situation gets worse, it is likely that the booking area will need to be expanded.

Does that sound familiar? It is the description of the situation that you read earlier. It is quite a good problem statement. The jail administrator was lucky: The first task was already completed.

Now that the jail administrator has a good problem statement, what is the next step?

When you have completed your answer, go to the next page.

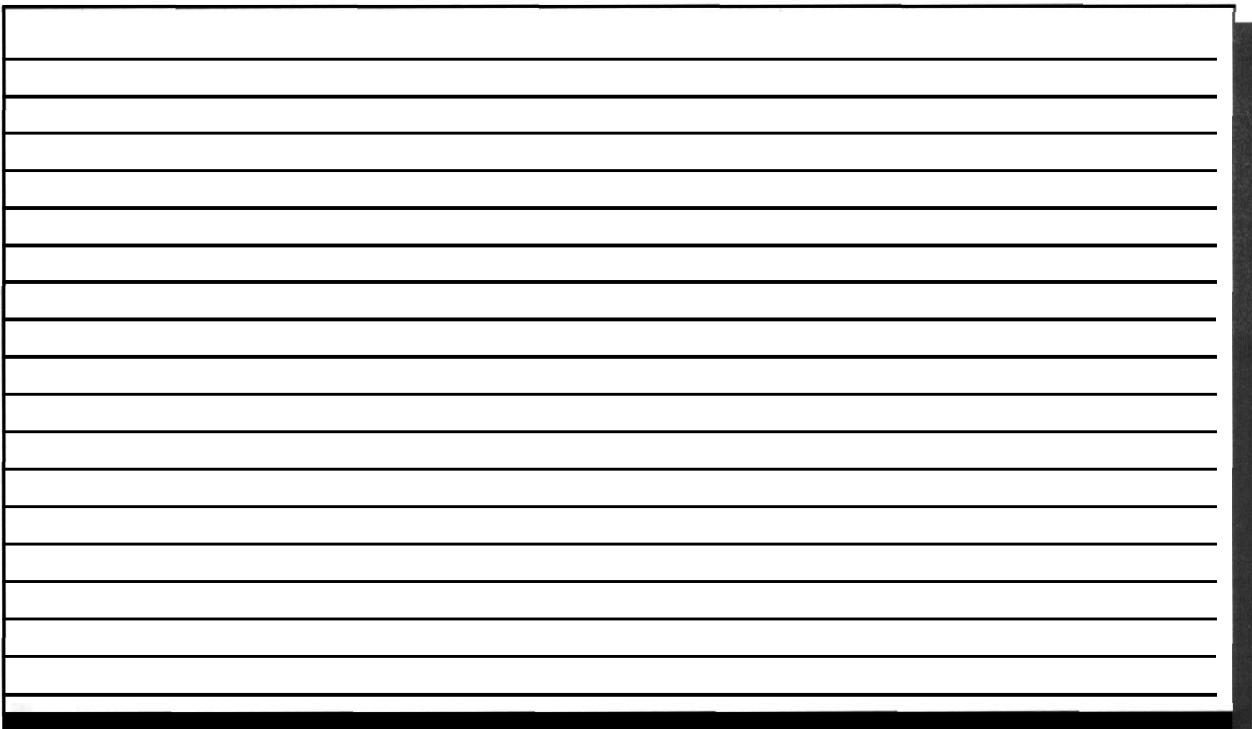
The jail administrator identified the questions that had to be answered in the analysis and the **data elements** needed. The jail administrator identified the following questions:

1. *How much have detox holds increased over the last 3 years?*
2. *Are the detox holds brought to the jail violent?*
3. *Are detox holds brought to the jail when there is space for them at the detox center?*
4. *Will the number of detox holds stabilize, decrease, or increase?*

The jail administrator identified the following data elements that were needed:

- *The number of detox holds detained at the jail by month as far back as possible,*
- *The number of detox holds who commit incidents in the Booking Room, and*
- *The number of times that the detox center was full during each year.*

What should the jail administrator do next?



When you have completed your answer, go to the next page.

The jail administrator was concerned about whether the jail kept information about detox holds since they were never booked. However, the jail clerk who maintained the information system indicated that a special category for detox holds had been established, records had been kept about detox holds for about 5 years, and the information system could provide the required historical data.

The jail administrator was also worried about identifying the number of detox holds who were violent. Jail records could provide the number of detox holds involved in incidents at the jail. It was the jail's policy that whenever someone had to be restrained or behaved in a violent way, the officer(s) involved wrote an Incident Report. The reports were filed in both an Incident Log and the inmate's record. The jail administrator would be able to review the Incident Log to identify how many detox holds were involved in incidents in the facility.

The question that was the hardest for the jail administrator to answer was how often people were brought to the jail because the detox center was full. The minutes of the task force meeting indicated that the detox center had been operating at about 50% of capacity, but the jail administrator was suspicious of averages. Worse, the jail had no way to gather that data.

The jail administrator did not know whether detox holds were likely to increase, decrease, or remain stable. He hoped that the available information would provide the answer.

Assume that the jail administrator got all the data together. He thought it might be a good idea to share this information with the sheriff first. This is wise. No one likes to be surprised about data from his/her own organization.

What information should the jail administrator prepare for the meeting with the sheriff?

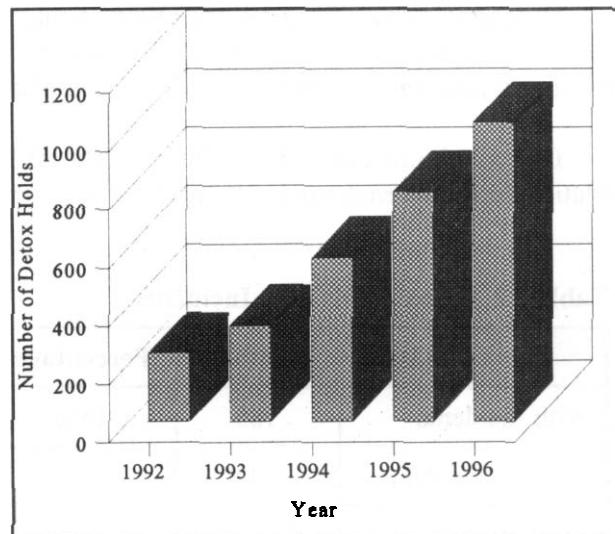
When you have completed your answer, go to the next page.

The jail administrator used both tables and charts to summarize the information he gathered. Table 7.1 and Figure 7.1 display the number of detox holds detained at the jail for the past 5 years.

Table 7.1 Historical Trend in Detox Holds

Year	Detox Holds
1992	237
1993	332
1994	564
1995	790
1996	1,027

Figure 7.1 Historical Trend in Detox Holds



What do the table and figure suggest?

[A large rectangular area with horizontal lines for writing.]

When you have completed your answer, go to the next page.

Both the sheriff and jail administrator were surprised to find that detox holds had increased 333% from 1992 to 1996. This increase made them nervous. Even if additional booking space was built, it would be at least 3 years until that space was available. And if detox holds increased so much in 4 years, things could get worse before new space was built.

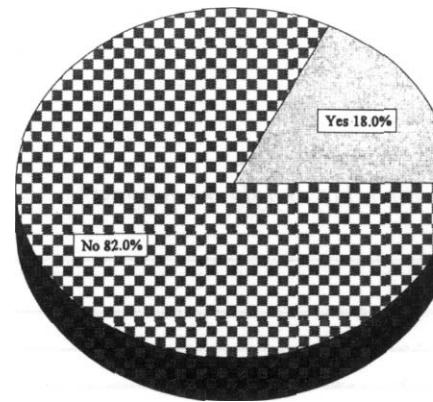
The sheriff knew of a technique called population forecasting, which allows statisticians to predict what a future population was likely to be. Recognizing that this was an advanced statistical technique, the sheriff asked a staff member, who was a graduate student at a local university, for assistance.

The jail administrator next showed the sheriff Table 7.2 and Figure 7.2, which displayed the relationship between detox holds and incidents committed in the jail in 1996.

Table 7.2 Detox Holds with Incidents

Detox Holds	Number	Percentage
With Incidents	185	18%
Without Incidents	842	82%
Total	1,027	100%

Figure 7.2 Detox Holds with Incidents



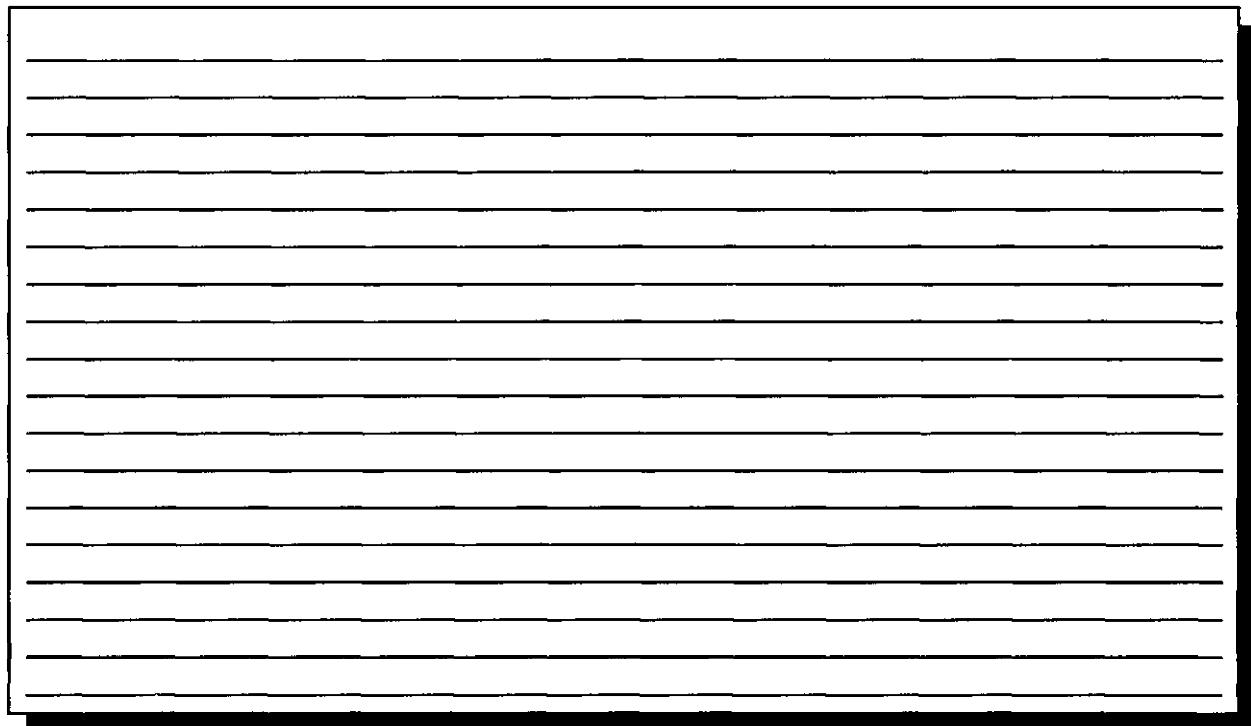
What do the table and figure suggest?

When you have completed **your** answer, go to the next page.

Both the sheriff and jail administrator were surprised by how few of the detox holds were involved in incidents while in the Booking Room. A review of the Incident Log Book revealed that, although some of the incidents that **did** involve detox holds were violent (usually because the person was violent when brought in), a number involved medical or health-related problems. Only 185 of the 1,027 detox holds that were brought in that year were violent.

The jail administrator thought that the odds were against the detox center being full the other 842 times that the jail had detained people. The jail administrator had talked with the detox center administrator, who indicated that the detox center was at capacity only about one day a month. Even if this provided a reason for some of the detox holds, it left a great many **not** accounted for.

What does this information suggest?

A large rectangular box with a black border, designed to look like a sheet of lined paper. It contains 20 evenly spaced horizontal lines for writing a response to the question "What does this information suggest?"

When you have completed your answer, go to the next page,

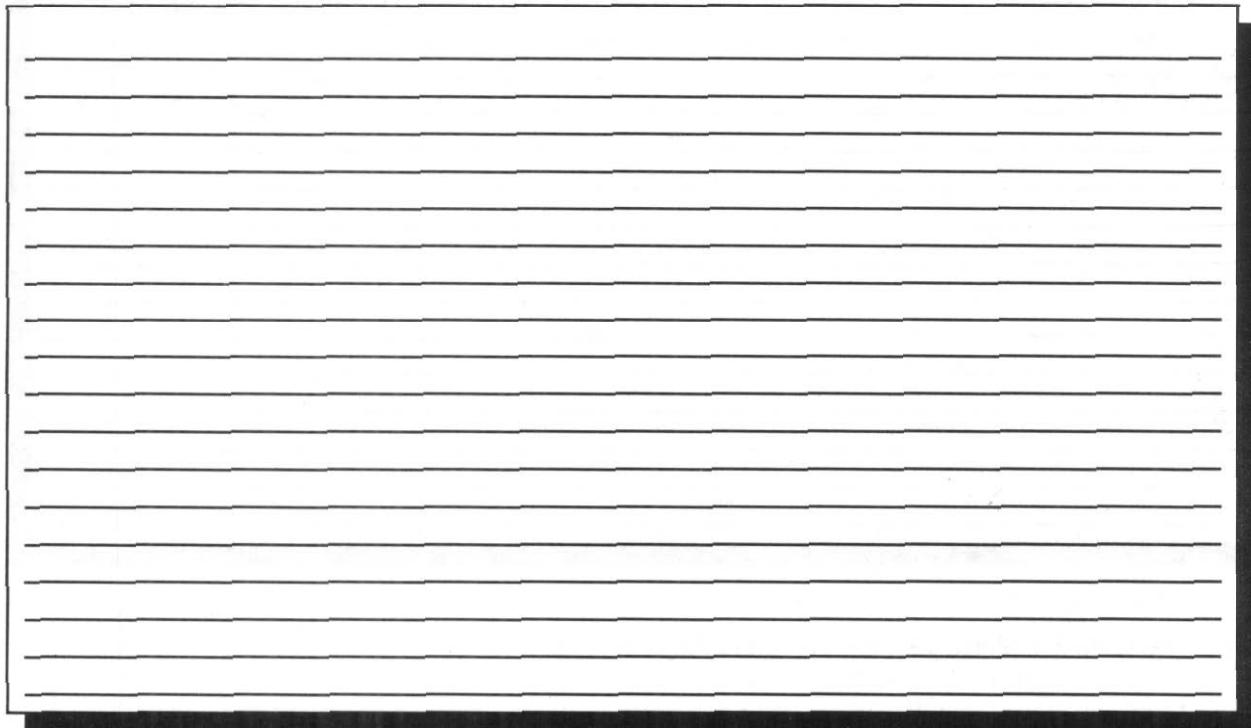
Both the sheriff and jail administrator concluded that a number of detox holds coming to the jail did not meet the criteria for detention. Remember...

The Metro County Jail may detain persons for the local detox center:

1. *If they are violent, or*
2. *If the detox center is full.*

While the sheriff and the jail administrator, both former patrol officers, were beginning to suspect what might be occurring, they did not want to jump to conclusions. They agreed that some additional information about detox holds and why they were brought to the jail was necessary. The task of obtaining that information was assigned to the jail administrator. The sheriff and jail administrator agreed to meet one week later to decide what action (if any) the jail should take and what they would recommend to the task force.

What did the jail administrator do next?



When you have completed your answer, go to the next page.

The jail administrator considered a number of different strategies, including:

1. Asking patrol officers about their past use of the detox center for intoxicated people, and
2. Identifying several dates when there were a number of detox holds, pulling the complaint reports for them, and calling the detox center to find out whether it was full on those dates.

The jail administrator, however, rejected both strategies. The patrol officers might not remember, or there would not be any documentation about why the detox holds had been brought to the jail. In addition, the thought of going through all that paperwork was discouraging. Instead, the jail administrator developed a third strategy. The following directive (Figure 7.3) for booking officers was written and read at all shift briefings that day.

Figure 7.3 Memo from the Jail Administrator

I N T E R _____ O F F I C E	MEMO
<p>To: All Booking Officers From: Jail Administrator Subject: Detox Holds Date: September 27, 1997</p> <p>Effective immediately, when a patrol officer brings in a detox hold, booking officers will:</p> <ol style="list-style-type: none">1. Ask the patrol officer why the detox hold is being brought to the jail rather than the detox center and make sure that this reason is recorded on the Booking Form.2. Accept the detox hold.3. If the detox hold is violent at the time of admission, attach a copy of the Incident Report to the Booking Form and forward to the jail administrator.4. If the patrol officer indicates that the detox hold was violent at the detox center, call the detox center for verification after the patrol officer has left, record this information on the Booking Form, and forward to the jail administrator.5. If the patrol officer indicates that the detox center was full, call the detox center for verification after the patrol officer has left, record this information on the Booking Form, and forward to the jail administrator.	

While the jail administrator was concerned about how patrol officers might view this strategy, he thought it was more important to avoid disagreements between the booking and patrol officers (particularly in

front of the detox hold) while the information was being gathered. He was also somewhat concerned about the size of his sample. The current week had only 25 detox holds.

Nevertheless, the jail administrator, thinking that some information is better than none, compiled the available information in a table(7.3) and figures(7.4 and 7.5) for the meeting with the sheriff.

Table 7.3 Patrol **Officers'** Reasons for Detox Holds

Patrol Officer Said:	Jail Finding	Number	Percentage
Violent at Detox Center	Yes	1	4%
	No	15	60%
Detox Center Full	Yes	2	8%
	No	7	28%
Total Detox Holds		25	100%

Figure 7.4 Detox Holds Violent at Detox Center

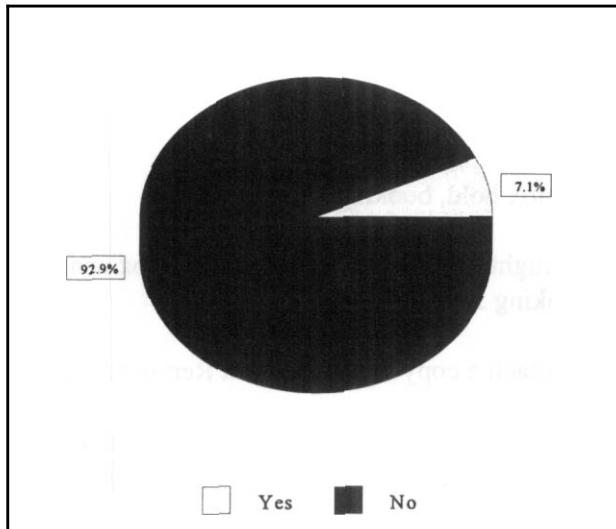
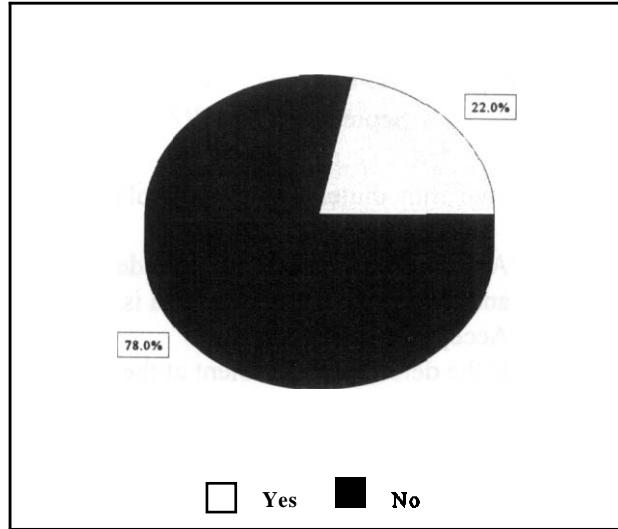
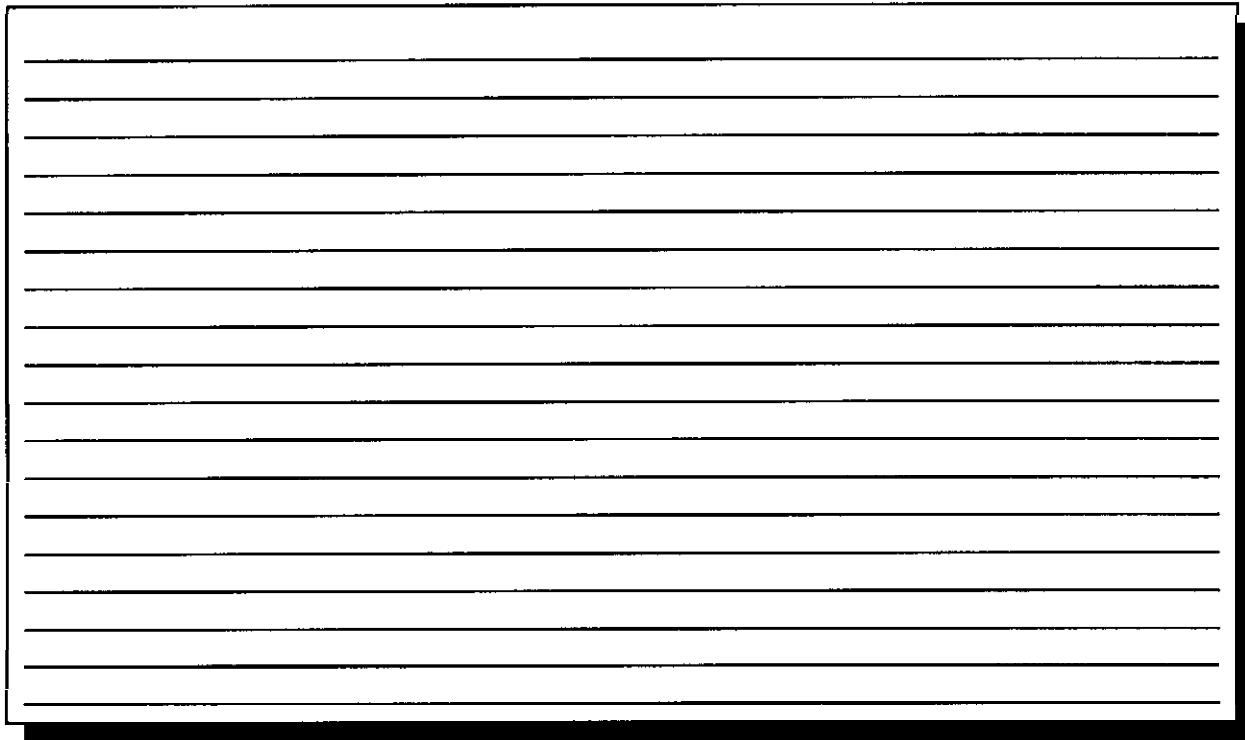


Figure 7.5 Detox Holds **When** Detox Center Full



What conclusion would you draw from this sample?



A large rectangular box with a black border, designed to look like a sheet of lined paper. Inside the box are 20 evenly spaced horizontal lines for handwriting practice.

When you have completed your answer, go to the next page,

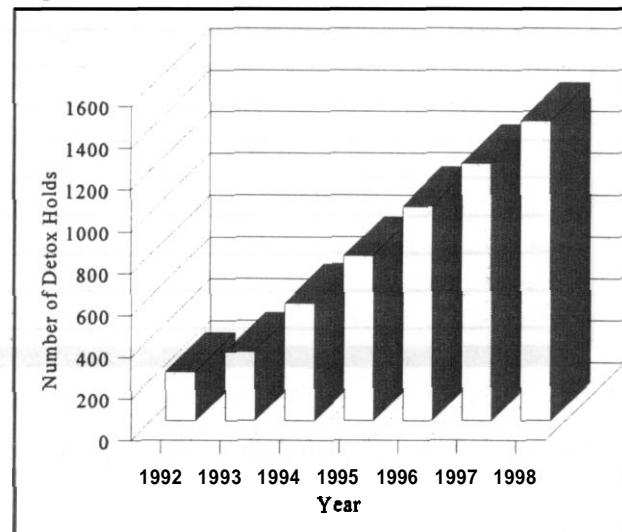
Both the sheriff and jail administrator concluded that patrol officers were bringing detox holds to the jail that should have gone to the detox center. The jail administrator was able to verify this—several of the booking officers had noted on the Booking Form that patrol officers who said they first took people to the detox center had never been there.

The sheriff then showed his projected increases in the number of detox holds that would be held at the jail. The analysis began by summarizing the number of detox holds for 5 years, although the forecast was actually made by using data grouped by month. The numbers indicated that if the present trend in detox holds detained at the jail continued, they could expect to have 1,231 (about 3.5 a day) in 1997 and 1,435 (about 4 a day) in 1998. See Table 7.4 and Figure 7.6.

Table 7.4 Trend in Detox Holds

Year	Detox Holds
1992	237
1993	332
1994	564
1995	790
1996	1,027
1997 est.	1,231
1998 est.	1,435

Figure 7.6 Trend in Detox Holds



Based on this information, what should the sheriff and jail administrator recommend to the task force?

When you have completed your answer, go to the next page.

A wide range of actions could be taken to deal with the large number of detox holds brought to the jail instead of the detox center. The sheriff knew that a city police department was the primary user of the jail (with about 75% of all inmates and nearly all detox holds). The first approach to solving the problem was to discuss it with the local police chief. Unfortunately, they were not able to come to an agreement.

Not successful in his attempt at collaboration, the sheriff decided that acting to preserve the jail's interest in this situation was consistent with local criminal justice system policy. The detox center was developed as an alternative to jail after public intoxication was decriminalized, to avoid having to increase the jail's capacity.

In his response to the task force, the sheriff indicated that the jail had detained detox holds as a favor to the local criminal justice system and the detox center, but that this favor had been somewhat abused. Detox holds had become a serious problem for the jail on weekends as they used bed space the jail needed for persons who had been arrested. While the current situation was bad enough, detox holds were projected to increase at such a rate that, if no changes were made, the Booking Room would require expansion within the next 2 years. The sheriff concluded that expanding the Booking Room to accommodate the detox holds would represent a major shift in criminal justice system policy.

The sheriff established a policy that the jail would no longer accept detox holds. Those holds who were violent could be charged with a crime. The detox center indicated that, although it was seldom full, it had the space to expand if necessary. This was not an easy decision to make—nor a very popular one with some of the local law enforcement community.

Six months later, the detox center reported that it was operating at close to 95% of capacity and was adding 10 more beds. It changed its weekend staffing pattern to more adequately manage the people placed in its custody. The jail reported that space problems in the booking area were much improved. Even though Friday and Saturday nights were still busy, the jail no longer had people sleeping it off for 8 hours in the three available holding cells while officers tried to book others charged with crimes.

Before Going on to Case Study Two

Clearly, statistics are used with both qualitative information (the sheriffs meeting with the police chief) and criminal justice system policy preferences (the system's choice to use the detox center as a way of dealing with alcohol problems) to make policy decisions. In the first case study, the sheriff made a policy decision that was both proactive and supportive of the jail's position. Other sheriffs might have made different but equally effective decisions. Without the information gathered in the study, the sheriff might have recommended to the task force that the Booking Room be expanded immediately, resulting in the expenditure of county funds to house people in the jail that the criminal justice system had determined should be housed elsewhere.

By now, the process of analysis should be clearer. Take a look at the second case study and see what decisions you make to resolve a problem many counties face when they plan new jails: Their needs cannot be accommodated by their resources.

Case Study Two: The Alternative Answer

The Situation

Forest County encompasses a relatively small, rugged area and has a population of about 65,000. The 60-bed jail, constructed in 1935, is crowded, with an ADP of 81.25 inmates. Because of a Federal District Court consent decree, the county has been preparing to build a new jail and sheriffs office for about a year. A task force¹ has been working with a criminal justice planning firm to complete a needs assessment. Population projections indicate that a jail of approximately 250 beds will be needed to meet anticipated needs until 2020. Unfortunately, the county can afford a facility of only 150 beds.

After the initial budgeting session, the task force decided to rethink a number of criminal justice policies. Task force members asked their planning consultant to help them determine potential segments of the jail population that might be appropriate for alternatives to incarceration. They identified a number of alternatives that would be appropriate for the community, such as house arrest with electronic monitoring, day reporting, boot camp, pre-trial case processing modifications, and intensive supervision.

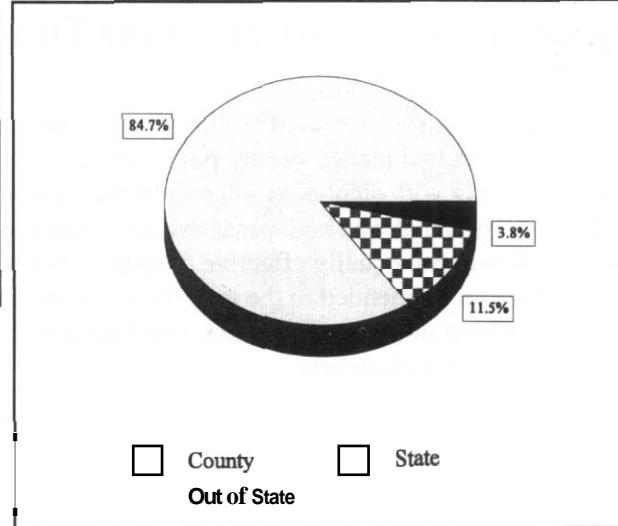
The criminal justice planner provided task force members with an Inmate Profile Data Collection Report, based on a sample of more than 600 cases collected over a 6-month period in 1998. The task force was asked to help identify the alternatives that would have the greatest impact on the jail population.

You are the sheriff. You begin to review the Inmate Profile Data Collection Report. Table 7.5 and Figure 7.7 catch your attention.

Table 7.5 Resident Status

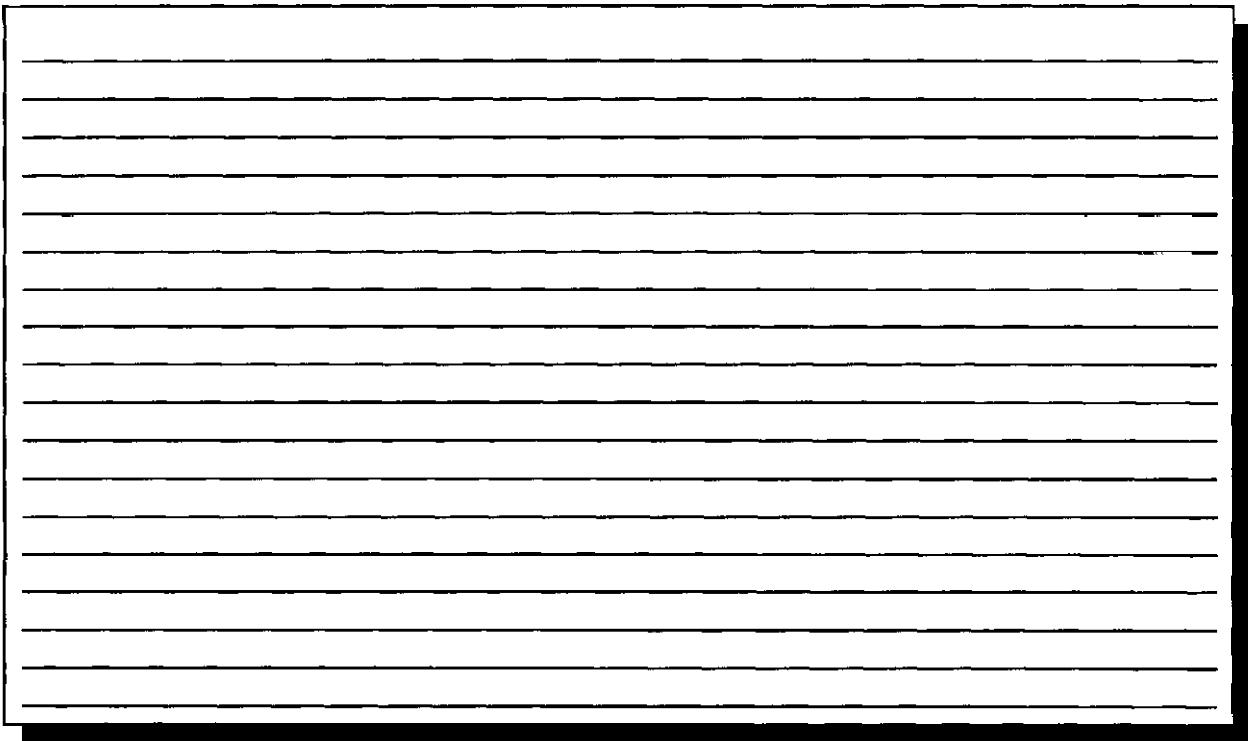
Resident Status	Number	Percentage
County	574	85%
State	78	11%
Out of State	26	4%
Total	678	100%

Figure 7.7 Resident Status



The Jail Task Force includes the sheriff, the chief of police of the primary city, the presiding judges of the county and district courts, the prosecutor, the chief probation officer, and two representatives of the five-member county commission.

How would you analyze and use this information?



A large rectangular area containing 20 horizontal lines for writing.

When you have completed your answer, go to the next page

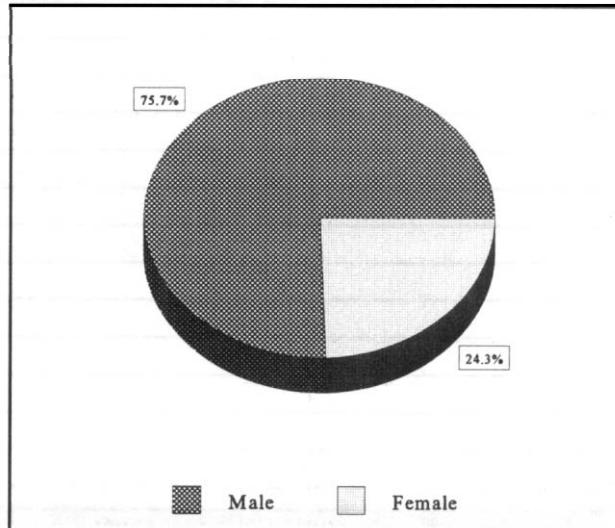
The sheriff found that most of the people detained in the jail were county residents. It would be easier to decide if they would be reasonable candidates for a sentencing alternative and they would be easier to find later.

The sheriff looked a little further. Table 7.6 and Figure 7.8 were next in the Inmate Profile Data Collection Report.

Table 7.6 Gender of Jail Residents

Gender	Number	Percentage
Male	513	76%
Female	165	24%
Total	678	100%

Figure 7.8 Gender of Jail Residents



What does this information suggest?

When you have completed your answer, go to the next page.

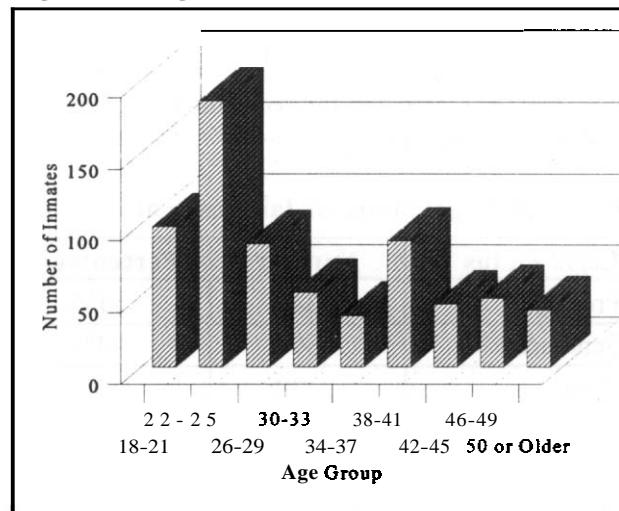
Table 7.6 troubled the sheriff because it seemed inconsistent with national trends. The sheriff went to a copy of the Bureau of Justice Statistics (BJS) Bulletin "Prison and Jail Inmates at Midyear 1996." This report indicated that women are about 11% of the national jail population. About 24% of the inmate admissions to the Forest County jail were women. Even if the BJS statistics dealt with **inmates in custody** and the sheriff was counting bookings, the difference still seemed too great. The sheriff wondered if this was an important clue to something unusual in the system's practices.

Table 7.7 and Figure 7.9 were next in the study.

Table 7.7 Age of Jail Residents

Age	Number	Percentage
18-21	98	14%
22-25	186	27%
26-29	86	13%
30-33	52	8%
34-37	36	5%
38-41	88	13%
42-45	44	6%
46-49	48	7%
50 or older	40	6%
Total	678	100%

Figure 7.9 Age of Jail Residents



What does this information suggest?

When you have completed your answer, go to the next page.

Table 7.7 did not seem out of the ordinary. While 45% of the jail population was over 30 (which may make it slightly older than the population in many jails), this was not very different from what the sheriff had expected. The sheriff suspected that it might not make much difference in this situation, because few alternatives or programs specifically target a particular age of offender.

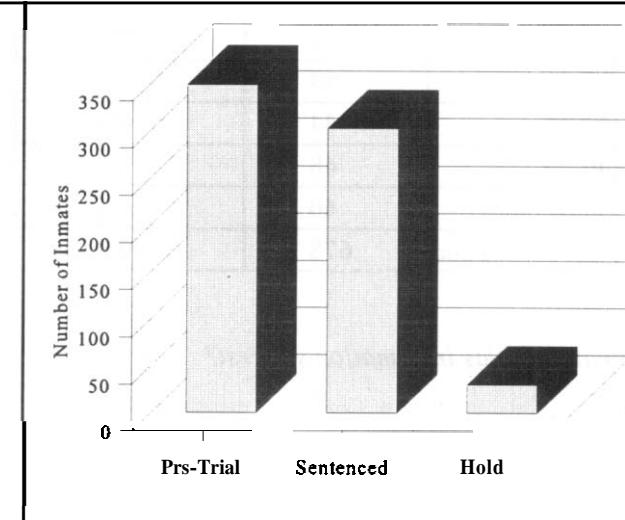
This table represents one of the problems in analyzing information: All the information available does not pertain to the problem. One of the critical tasks in interpreting the results of data collections is to sort out the information that does make a difference from all the information generated. Computers, in conjunction with poorly focused analysis, frequently cause "information overload," in which too much information is processed and presented for the analysts to comprehend. The sheriff had learned to sort out the things that were unusual or unexpected.

The next table and figure in the Inmate Profile Data Collection were about Legal Status. See Table 7.8 and Figure 7.10.

Table 7.8 Legal Status of Jail Residents

Legal Status	Number	Percentage
Pre-Trial	347	51%
Sentenced	301	44%
Hold	30	4%
Total	678	100%

Figure 7.10 Legal Status of Jail Residents



What does this information suggest?

When you have completed your answer, go to the next page.

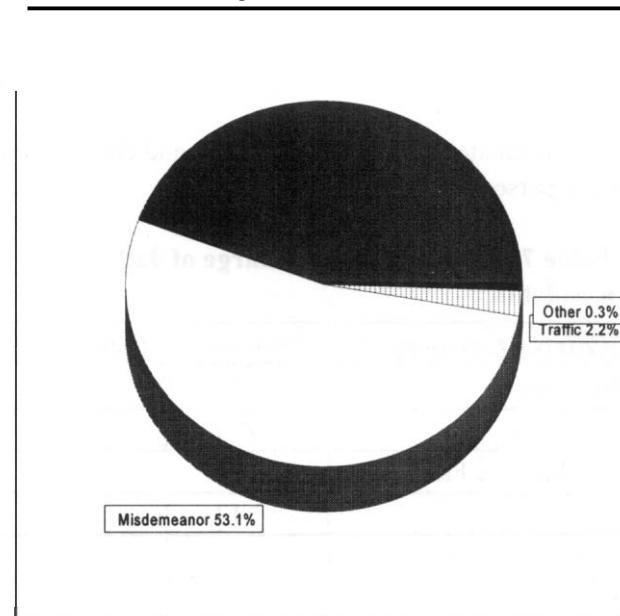
Table 7.8 indicated that 51% of the people booked at the jail were pre-trial detainees and 44% were sentenced. The remainder were held only for another jurisdiction on a warrant. Any alternatives that the jail would use to manage the jail population needed to work with both populations.

The next table and chart provided information about charge level.

Table 7.9 Charge Level of Jail Residents

Charge Level	Number	Percentage
Felony	301	44.40%
Misdemeanor	360	53.10%
Traffic	15	2.21%
Other	2	0.29%
Total	678	100.00%

Figure 7.11 Charge Level of Jail Residents



In deciding which alternatives will be most effective, how do you interpret this information and what else does it make you want to know?

[A large rectangular area with horizontal lines for writing.]

When you have completed your answer, go to the next page.

The sheriff was not surprised by the data shown in Table 7.9. Just under half of the population were charged with felonies—the hardest inmate group to divert from jail. Fifty-five percent were charged with either misdemeanors or traffic offenses **carrying** the weight of a misdemeanor. Charge certainly was a major consideration in matching inmates with alternatives. The sheriff knew that placing certain types of inmates in alternative programs would be like waving a red flag in the faces of a relatively conservative constituency. And it was clear that alternatives appropriate for a **pre-trial** misdemeanant would not work for a sentenced misdemeanant. The sheriff decided two additional questions must be answered:

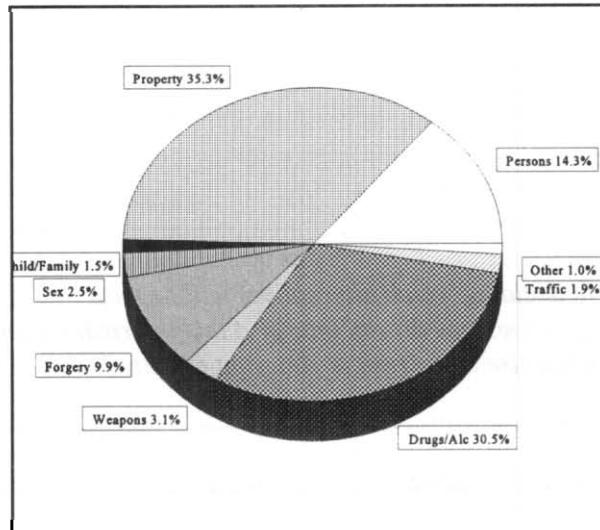
1. What were these inmates charged with?
 2. What was the relationship between charge level and legal status?

Fortunately, the following table and chart provided information about the type of offense for which each person had been booked.

Table 7.10 Most Serious Charge of Jail Residents

Offense Category	Number	Percentage
Persons	97	14%
Property	239	35%
Children & Family	10	1%
Sex	17	3%
Forgery & Fraud	67	10%
Weapons	21	3%
Drugs & Alcohol	207	31%
Traffic	13	2%
All Other	7	1%
Total	678	100%

Figure 7.12 Most Serious Charge of Jail Residents

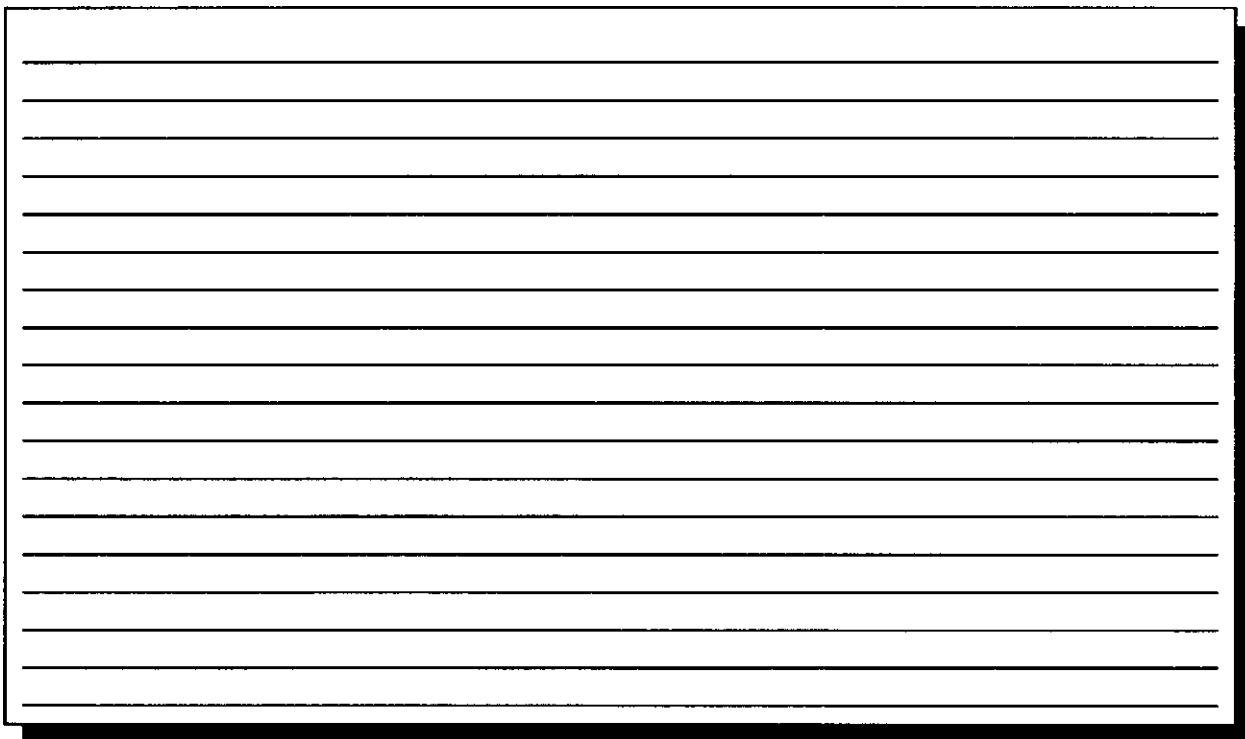


What are the implications of this information for alternatives to incarceration?

When you have completed your answer, go to the next page.

The sheriff noted that 14% of those who were booked in the jail were charged with offenses against persons. These were among the least likely candidates for diversion. Also, very few traffic offenders were booked, suggesting that the arrest standards calling for summons and citation of traffic offenders were being followed. The largest groups of inmates were booked for property and drug and alcohol offenses. The sheriff thought the task force might need some additional information about these inmates to determine how alternatives might be used with this part of the jail population.

What other things would you want to know about the property and drug and alcohol offenders?



A large rectangular box with a black border, designed for writing. It contains 20 horizontal lines spaced evenly apart, intended for the user to respond to the question "What other things would you want to know about the property and drug and alcohol offenders?"

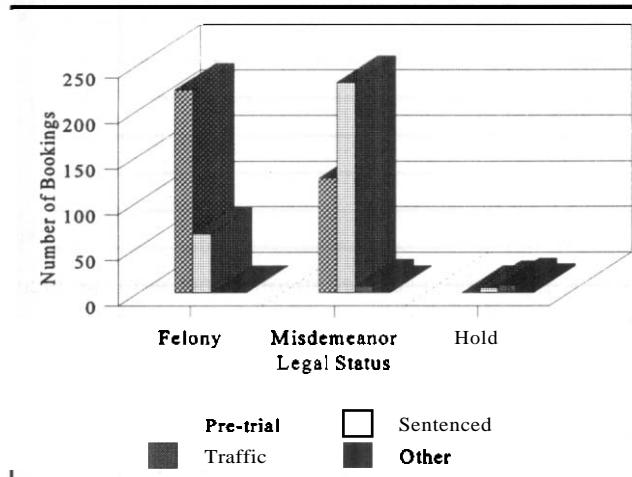
When you have completed your answer, go to the next page

The sheriff then examined the report for information about the relationship between legal status and charge level. The sheriff found a cross-tabulation of the two data elements in Table 7.11 and Figure 7.13.

Table 7.11 Relationship Between Legal Status and Charge Level

Charge Level	Legal Status					Total	
	Pre-Trial		Sentenced		Hold		
Felony	222	33%	64	9%	15	2%	301
Misdemeanor	125	18%	230	34%	5	1%	360
Traffic	0	0%	7	1%	8	1%	15
Other	0	0%	0	0%	2	0%	2
Total	347	51%	301	44%	30	4%	678

Figure 7.13 Legal Status by Charge Level



What are the implications of the cross-tabulation of charge level and legal status?

When you have completed your answer, go to the next page.

The sheriff was feeling a certain amount of frustration with this business of "analysis." It seemed that the more that you learn, the more you need (or want) to know. The sheriff decided that the criminal justice planner should have answers to the following questions at the next task force meeting:

1. What were the **specific** charges in the property and drug and alcohol categories?
 2. How many people had been booked for each?
 3. If this was the Most Serious Charge, how many **other** charges were involved:)
 4. Were any of those charges crimes against persons?
 5. Did the person have any outstanding warrants from other jurisdictions?
 6. Was there a Probation or Parole Violation in addition to this charge?

The implications of the cross-tabulation seemed clear to the sheriff. Three large groups (subpopulations in statistical language) were present in the jail. They were:

1. Pre-trial felons,
 2. Pre-trial misdemeanants, and
 3. Sentenced misdemeanants.

The sheriff talked with the jail administrator about these three groups. Developing alternatives for pre-trial felons was likely to be challenging, and alternatives for sentenced misdemeanants had complications of their own. However, pre-trial misdemeanants seemed to be a relatively easy category to find alternatives for. A good review of the bond schedule or a pre-trial release program should do the trick. The jail administrator, however, was puzzled by this data, which seemed inconsistent with the faces that she saw every week during the facility inspection. And yet, the sheriff and jail administrator knew that the information had come directly from their automated booking system. They contacted the criminal justice planner who had analyzed the data to see if this apparent discrepancy could be resolved.

What could cause this discrepancy?

When you have completed your answer, go to the next page.

The criminal justice planner suggested to the sheriff and jail administrator that it was possible they were dealing with a common jail phenomenon: the difference between the number of people booked and the number of people in custody on any given day. The easiest way to illustrate this was by the following tables, which were to be part of the final report to the task force.

Table 7.12 Persons Booked for Property and Drug and Alcohol Offenses

Bookings	Pre-Trial		Sentenced		Total
	Felony	Misdemeanor	Felony	Misdemeanor	
Burglary	33	0	9	0	42
Criminal Mischief	3	11	6	35	55
Theft	37	47	14	28	126
All Other Property	0	0	3	13	16
Subtotal Property	73	58	32	76	239
Possession	34	14	6	37	91
Sales	21	0	12	0	33
DUI	1	8	7	67	83
Subtotal Drugs & Alcohol	56	22	25	104	207
Total	129	80	57	180	446

Table 7.13 Average Daily Population of Property and Drug and Alcohol Inmates

Average Daily Population	Pre-Trial		Sentenced		Total
	Felony	Misdemeanor	Felony	Misdemeanor	
Burglary	3.35	0.00	4.44	0.00	7.78
Criminal Mischief	0.11	0.06	1.48	2.88	4.52
Theft	1.72	0.39	3.45	2.30	7.86
All Other Property	0.00	0.00	0.37	1.07	1.44
Subtotal Property	5.18	0.45	9.74	6.25	21.61
Possession	1.30	0.08	2.96	4.05	8.39
Sales	1.21	0.00	6.12	0.00	7.32
DUI	0.01	0.02	3.45	5.51	8.99
Subtotal Drugs & Alcohol	7.69	0.55	22.27	15.81	46.31
Total	12.87	0.99	32.01	22.05	67.92

The jail administrator and sheriff looked at Tables 7.12 and 7.13 and said, "So what?" Fortunately, Tables 7.14 and 7.15 made the criminal justice planner's point clearer.

Table 7.14 Percent of Persons Booked for Property and Drug and Alcohol Offenses

Bookings	Pre-Trial		Sentenced		Total
	Felony	Misdemeanor	Felony	Misdemeanor	
Offense Category					
Burglary	7%	0%	2%	0%	9%
Criminal Mischief	1%	2%	1%	8%	12%
Theft	8%	11%	3%	6%	28%
All Other Property	0%	0%	1%	3%	4%
Subtotal Property	16%	13%	7%	17%	54%
Possession	8%	3%	1%	8%	20%
Sales	5%	0%	3%	0%	7%
DUI	0%	2%	2%	15%	19%
Subtotal Drugs & Alcohol	13%	5%	6%	23%	46%
Total	29%	18%	13%	40%	100%

Table 7.15 Percent of Average Daily Population of Property and Drug and Alcohol Inmates

Average Daily Population	Pre-Trial		Sentenced		Total
	Felony	Misdemeanor	Felony	Misdemeanor	
Offense Category					
Burglary	5%	0%	7%	0%	11%
Criminal Mischief	0%	0%	2%	4%	7%
Theft	3%	1%	5%	3%	12%
All Other Property	0%	0%	1%	2%	2%
Subtotal Property	8%	1%	14%	9%	32%
Possession	2%	0%	4%	6%	12%
Sales	2%	0%	9%	0%	11%
DUI	0%	0%	5%	8%	13%
Subtotal Drugs & Alcohol	11%	1%	33%	23%	68%
Total	19%	1%	47%	32%	100%

*What conclusions can be drawn from **these** tables?*

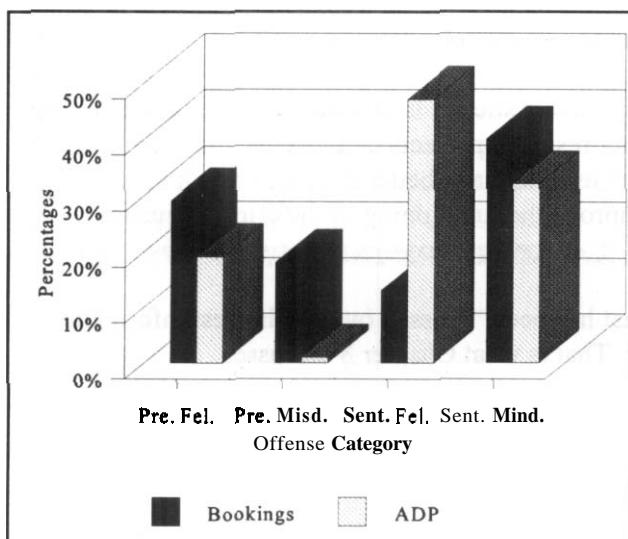
A large rectangular area with horizontal lines for writing, framed by a thick black border.

When you have completed your answer, go to the next page.

The sheriff and jail administrator noticed substantial differences between the first two tables when they were expressed as percentages in the second two tables.

1. Although pre-trial felons in the two categories reviewed were 29% of the bookings, they accounted for only 19% of the ADP of the facility.
2. Pre-trial misdemeanants were 18% of the bookings in the two categories, but only 1% of the ADP.
3. Sentenced felons were 13% of bookings, but 47% of the ADP.
4. Felons sentenced for drug and alcohol offenses were 6% of bookings, but 33% of the ADP.
5. Sentenced misdemeanants were 40% of bookings in the two categories, but 32% of the ADP

Figure 7.14 Comparison of Bookings and ADP of Property and Drug and Alcohol Offenders



The average Length of Stay for sentenced felons in these categories was 142.5 days. If their average Length of Stay could be reduced to 90 days, the ADP of Sentenced Felons would be reduced from 32 to 14.

The sheriff and jail administrator saw a way through what the task force feared would be a difficult situation politically. No one was going to divert sentenced felons from the jail. Part of their sentence would be served in the jail, but the remainder of it could be served in any of the nonresidential alternatives. The sheriff and jail administrator developed a proposal for electronic home detention, day reporting, and intensive supervision as part of the sentence. The task force considered this proposal an appropriate alternative for the county and recommended its implementation.

Figure 7.14 made the point even clearer. In examining the groups that populated the jail, it was critical that the task force members think in terms of ADP, Inmate Days, or another measure of how much of the jail bed space a group of inmates actually uses. Clearly, different groups of inmates have very different Lengths of Stay. In making decisions about alternatives and how they might be used to "save" or "defer" jail bed space, it is a matter of both how many people come into the jail and how long they stay. Understanding this up front helped the task force members avoid a common problem—focusing on the wrong target group. For example, 18% of the bookings were pre-trial misdemeanants. The sheriff and jail administrator had considered that this might be a relatively easy group to divert from the facility. However, even if all of them disappeared, the ADP of the jail would have decreased only about 1%. However, while sentenced felons are only 13% of the bookings, they use nearly half of the jail bed space.

Conclusion

These case studies represent what it is really like to analyze data. The process is not as straightforward as statisticians may tell you and requires a measure of qualitative as well as quantitative information. Some areas are open to interpretation. Data collections often raise as many questions as they answer. You analyze information, discover you want to know something else, track down that information, reanalyze, and then reinterpret. Data analysis is much like good detective work. The "tools of the trade" may be different, but the thought processes are similar. Finally, analysis is very much about putting humans, not computers, in the role of interpreting what data means.

Interpreting jail or criminal justice system data is best done by people who know something about how the system works. Good analysts can **present** the data, identify things that seem unusual, locate targets for change, and perhaps point out potential alternatives, but they cannot make decisions. Sheriffs and jail administrators (and their counterparts in other parts of the criminal justice system) must take an active role in deciding what the information gathered means. Unless the managers in the system take part in the analysis, reports are shelved no matter how accurate or well done they are.

Sometimes, while collecting the data or calculating the statistics, it is possible to forget **why** they are being done. The data collections discussed here were not research projects or academic excursions into a practitioner's world. They were tools that real managers used to make better decisions about their organizations. They also provided an opportunity to improve the functioning of the criminal justice system.

There is one more critical part to data collection that has been "glossed over." The best information can still be overlooked if it is not displayed effectively. That is what Chapter 8 discusses.

Chapter 8: Sharing Information with Others

The final step in collecting and analyzing data is to share the results with others. The true test of a good analysis is if the data can be understood and used by someone else to solve problems or make decisions. Sometimes, analysts make it difficult for even the most determined people to use the data because information may be buried in too many figures or displayed ineffectively.

Much of the material in this chapter will be second nature to those who have been exposed to any of the common computer graphics programs or used the graphic functions of off-the-shelf spreadsheet and word processing software. However, good information often is displayed so poorly that it is discounted or disregarded because it is too difficult to interpret. This manual has provided methods by which sheriffs and jail administrators can improve their data collection and analysis. The data generated after such an effort merits the time and attention required to display it clearly and effectively. This chapter provides help in this area by:

1. Explaining why displaying data is important,
2. Describing and illustrating methods that can be used to display data,
3. Offering guidelines for displaying data, and
4. Showing "good" and "bad" examples of data that has been displayed.

Why Pictures Are Worth a Thousand Numbers

As Chapter 6 suggested, statistics have two results:

1. A **number(s)** that summarizes the information and
2. A graphic representation of the data.

Next, Chapter 7 focused on interpreting the **numbers**, as if the "pictures" in the data were not an important tool for interpretation. That may be backwards, because displaying the data is often an essential part of the analysis. Sometimes, the only thing that makes the data clear is to plot it, point by point. Displaying data effectively is critical for a number of other reasons.

1. Most people do not think in numbers. The non-mathematically minded person cannot conceptualize what "an increase of 937 net bookings over the previous year" really means. Pictures provide a way of making that point very clear.
2. When people try to remember things, they often attempt to visualize what it is they want to remember. Well-displayed data provides just such an image.
3. Carefully displayed data can highlight the critical points of the analysis. A well-constructed graphic will draw the eye of the viewer to exactly what analysts want him/her to remember.

4. Rows and rows of numbers are very boring, unless the audience is an atypical group that finds numbers more interesting than words or images. Well-displayed data can make reading about data more interesting.

Methods for Displaying Data

Statistical jargon sometimes makes it difficult to figure out what is being said about displaying data. Analysts talk about frequency polygons, histograms, cumulative frequency graphs, and other complicated terms for some very simple concepts.

There are four basic tools for displaying virtually all data. Computer graphics technology can enhance them and apply a variety of different "styles" to each.

1. Tables, in which the numbers are placed in columns and rows that have titles;
2. Bar charts, in which the height, order, and sometimes the width of the bars represent the size of one value or group in the data;
3. Pie charts, in which the size of a segment represents the number of cases in one of the categories of the data element; and
4. Line graphs, in which a line connects all related data points in a series.

These descriptions actually make a good case for using graphic examples, because the words are not as clear as a picture. An illustration of each method of display is given in each of the following discussions.

Tables

Tables display data by placing the numbers in titled columns and rows. Table 8.1 is an example of a "tabular display of data." Tables are used when numerical information is listed. Tables are useful ways to store data so that it will be easy to use again. Generally, tables clearly show what the actual numbers are. Although other figures may better show the relationships between categories of the data or two data elements, they are harder to read. As a result, most other figures require tables to clarify and present the exact numbers.

Table 8.1 Transports

Reason for Transport	Number	Percentage
Court	152	64%
Medical	40	17%
Warrant Pick-up	26	11%
Mental Health	12	5%
All Other	7	3%
Total	237	100%

The following steps improve the appearance of most tables:

1. Label the columns and the rows. The data displayed may be clear to you but a mystery to anyone else.
2. Make the table easy to read. Most word processing software has a "tables" function that puts data (either directly from a spreadsheet or manually entered) in boxes or "cells," like the examples in the tables in this document. The "tables" function has the advantage of separating each piece of information. The lines help your eye follow information across the table. If your word processing

software does not have a tables function, be sure to separate columns and rows by manually drawing in lines.

3. Do not put all the data into one table. The results look overwhelming. Most people tend to "skim" the tables in a report. This tendency can be overcome by making the tables short, neat, and to the point.
4. If the table displays all the categories of a data element (like Table 8.1), total the columns. If the table is a cross-tabulation (a frequency distribution of two data elements), total both the columns and the rows.

Tables help to organize data, but they do not create an image of the data. The following techniques do that.

Bar Charts

In bar charts, the height, order, and sometimes the width of the bars represent the size of one value or group in the data. Figure 8.1 takes the same data from Table 8.1 and displays it in a bar chart.

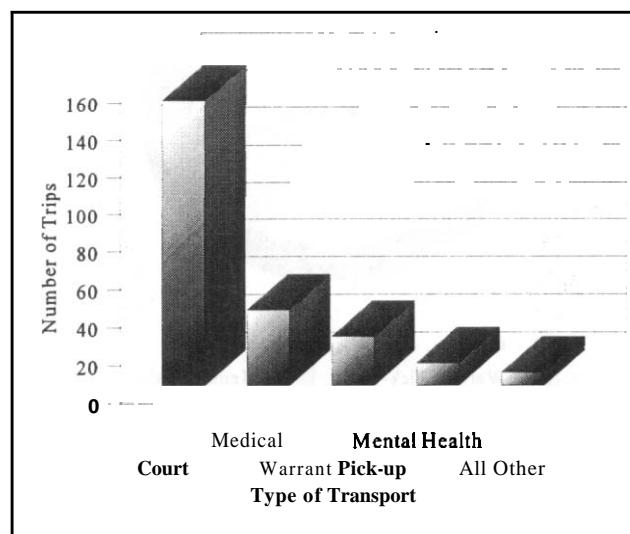
Bar charts are one of the most commonly used methods to display data because:

1. They are easy to read and understand.
2. They can make the major point of the data clear.
3. They are easy to construct.

Bar charts can be used with data at any level of measurement by following a few simple rules:

1. With nominal data (data elements that are divided into categories, like the data displayed in Figure 8.1), the bars should not touch. Only the relative heights of the bars are important.
2. With ordinal data (data elements that have groups ranked in some kind of order, but not measured in equal units), the bars should be in order. The heights have meaning.
3. With interval/ratio data (data elements that are measured in equal units), the area inside the bars should be proportional to the number of cases in each group (which affects the height and the width of the bar), and the bars should touch. This bar chart actually resembles the real-life frequency distribution of the data.

Figure 8.1 Reason for Transport



With practice, either with or without a computer graphics program, making a bar chart will be simple. Charts and graphs are oriented around two straight lines (each one called an "axis"), one horizontal (X axis) and one vertical (Y axis). Each line represents something. In Figure 8.1, the X axis tells what the

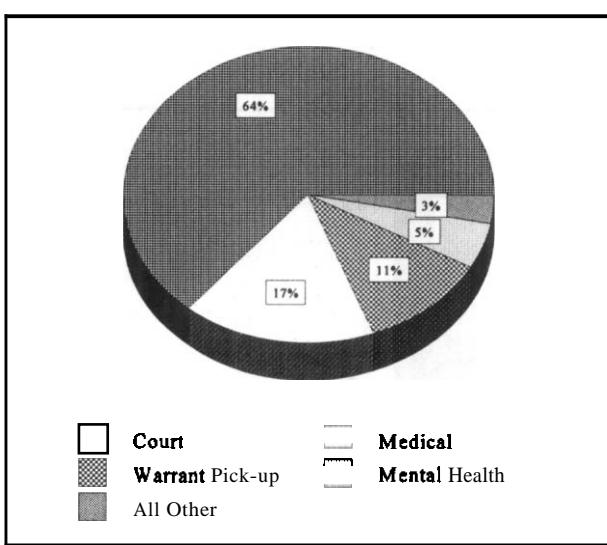
categories of the data element are and the Y axis tells how many cases, or what percent, are in each category. The Y axis must be divided into equal units. In other words, a "scale" must be developed for the Y axis. In graphics programs, computers automatically calculate the scale that is best to display the data. If you are not using a computer, Appendix K provides advice and hints.

Pie Charts

In pie charts, the size of the "piece of the pie" represents the number of cases in one of the categories of the data element. Pie charts can be used only with data grouped into categories and totaling 100%. Figure 8.2 is based on the same transport data displayed in Table 8.1 and Figure 8.1.

The ease of making pie charts with computer graphics programs is something of a two-edged sword. For example, just because you can make a pie chart that has 15 categories does not mean that people will be able to understand it. A general rule is to limit the number of slices to five. With more categories, combine the smallest ones into a category called "Other," then do a separate pie chart for the "Other" category.

Figure 8.2 Reason for Transport



Computer programs can allow you to use pie charts for comparisons by linking one or more. If you want their size to mean something, you can make them proportional. Table 8.2 and Figure 8.3 compare the transport data used previously with transport data from another year. Those not using automation should see Appendix K for hints on making pie charts.

Figure 8.3 Comparison of Transports in 1994 and 1995

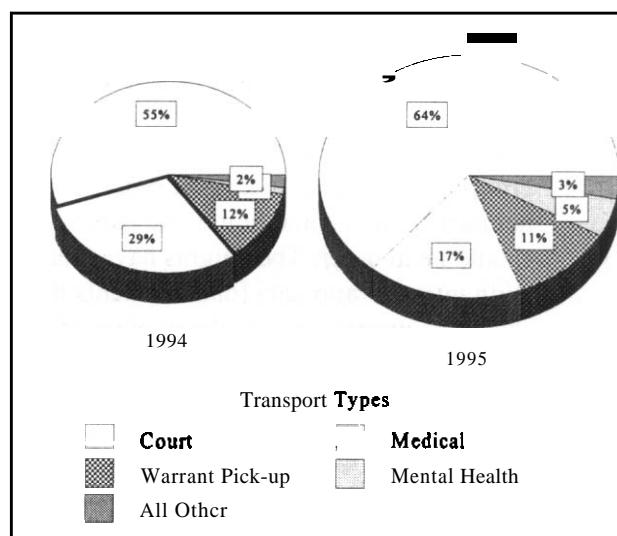


Table 8.2 Comparison of Transports in 1994 and 1995

Reason for Transport	1994	1995
Court	102	152
Medical	54	40
Warrant Pick-up	22	26
Mental Health	2	12
All Other	4	7
Total	184	237

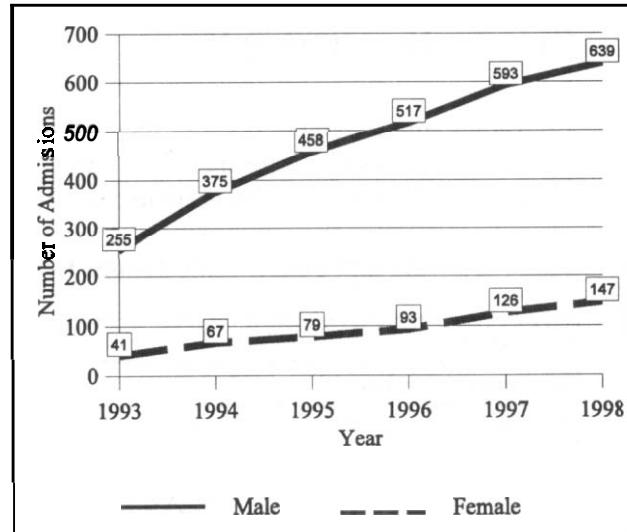
Line Graphs

Line graphs display data with a single line that connects all related data points in a series. Line graphs are common ways of displaying data. Table 8.3 represents male and female felony admissions from 1993 to 1998. Figure 8.4 illustrates these elements displayed on a line graph.

Table 8.3 Male and Female Felony Admissions (1993–1998)

Year	Male	Female
1993	255	41
1994	375	67
1995	458	79
1996	517	93
1997	593	126
1998	639	147

Figure 8.4 Male and Female Felony Admissions (1993–1998)



Line graphs can reveal a great deal about the distribution of the data if they are done properly. They should only be used for *interval/ratio* level data. They are easy to make with contemporary computer technology. Those who will make these graphs by hand should see Appendix K for helpful hints.

Guidelines for Good Graphics

The art of displaying data has "tricks of the trade." After all, an entire profession is based on the effective display of information.

Rule 1: Know your limitations.

Some people have more talent in this area than others. They **know** what looks good, what is "balanced." If you do not have these skills, delegate this task to someone who does. Individuals with good clerical skills and an "eye for detail" may be talented in these areas.

Rule 2: Make the graphic illustrate the most important information.

Analysis and interpretation of the data should make the most important point(s) clear. Display that information clearly. If the point is that since 1993 the number of felony admissions has risen steadily, use an example like Figure 8.4 to make that point clearly. Modern computer technology has given us a corollary to this rule.

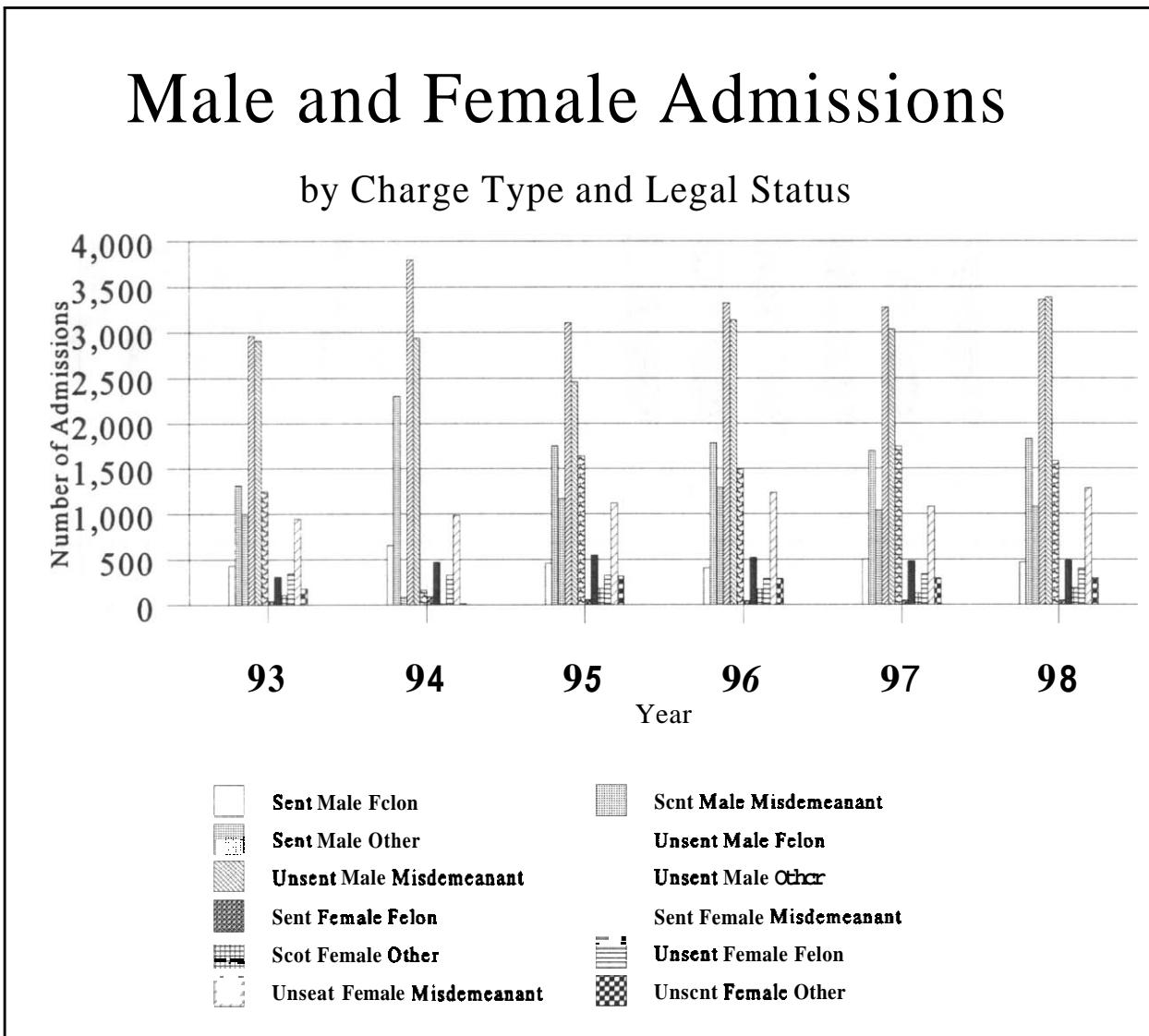
Rule 2a: Keep it simple.

Tables are convenient because they let you store much information compactly (such as Table 8.4, which contains a lot of information). Computer-generated graphics have little trouble handling this much information. Figure 8.5 is the graphic depiction of Table 8.4.

Table 8.4 Male and Female Admissions by Charge Type and Legal Status by Year

Category	1993	1994	1995	1996	1997	1998
Sentenced Males						
Felons						
Felons	433	655	458	405	503	469
Misdemeanants	1,317	2,304	1,751	1,782	1,694	1,826
Other	979	84	1,166	1,290	1,037	1,077
Unsentenced Males						
Felons						
Felons	2,961	3,799	3,105	3,319	3,270	3,355
Misdemeanants	2,912	2,939	2,461	3,133	3,033	3,382
Other	1,246	160	1,637	1,496	1,743	1,582
Sentenced Females						
Felons						
Felons	41	91	57	43	46	47
Misdemeanants	308	472	548	521	482	489
Other	106	11	177	171	127	180
Unsentenced Females						
Felons						
Felons	344	327	324	287	337	400
Misdemeanants	943	982	1,118	1,236	1,077	1,282
Other	177	14	317	290	291	293
Total	11,767	11,838	13,119	13,973	13,640	14,382

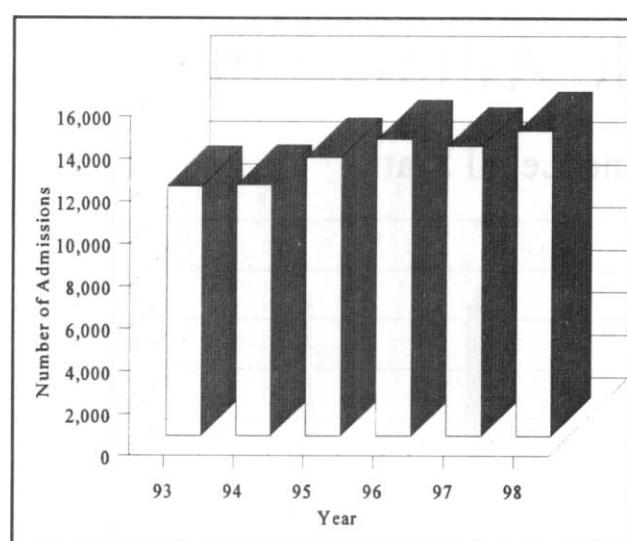
Figure 8.5 Male and Female Admissions by Charge Type and Legal Status by Year



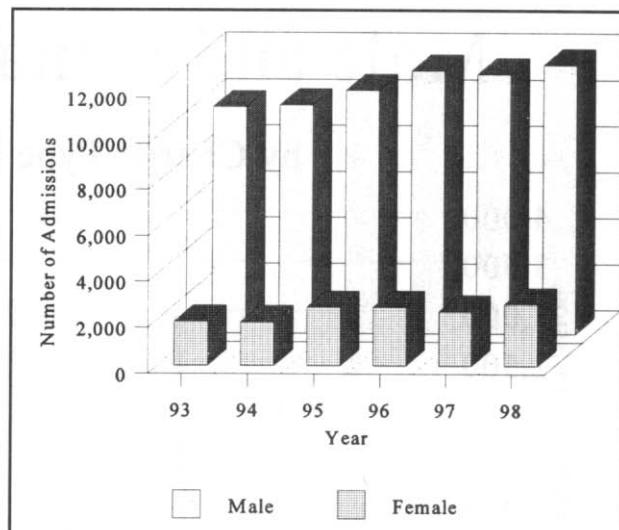
The computer may not have trouble with this, but the human eye does, as there are too many patterns to differentiate. And this figure must be larger to make it possible to even see the lines.

With this much information, you have a number of options. If your **primary** interest is how much admissions have increased, a simpler display is shown in Figure 8.6.

Figure 8.6 Admissions (1993–1998)



**Figure 8.7
Male and Female Admissions (1993–1998)**

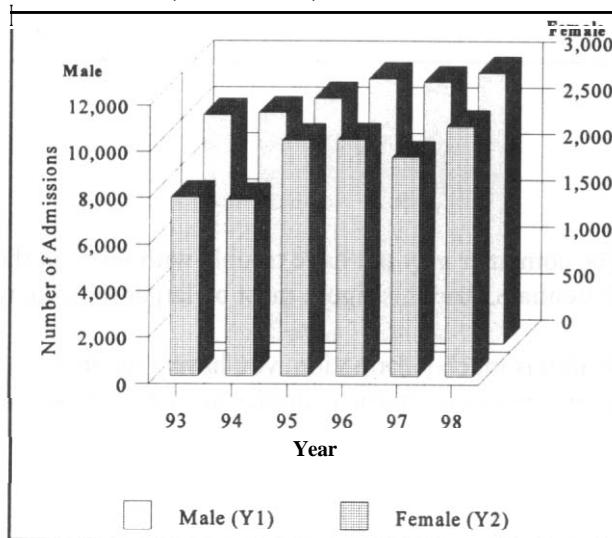


Sometimes less is more. Figure 8.6 is easier to read and shows the overall trend. You may show more detail about this data by focusing on a segment of the information. Figure 8.7 compares the trend in male and female admissions. Figure 8.8 takes the same information and examines the pattern of male and female admissions.

The difference between this figure and the previous one is that Figure 8.8 displays the data on two "Y axes" instead of **one**, as if the numbers were expressed as percentages of the total in each category. That lets the viewer examine the pattern within each **category**. For example, Figure 8.8 shows that, while the pattern of male and female admissions is quite similar for 1996 through 1998, female admissions increased markedly from 1994 to 1995.

If you really are interested in more detail in each of the categories, consider an option like Figures 8.9 and 8.10. Even these, which separate male and female, are very busy. So, before you **turn** your graphics program loose, think about the message in the data and plan a graphic to make that message most clear.

Figure 8.8 Male and Female Admissions (1993–1998)



Rule 3: Decide what is the best medium for sharing the information.

Most sheriffs and jail administrators tell other people about data in one of three ways:

1. A report or other written document,
2. A small working meeting, or
3. A large public meeting.

Information can be shared by:

1. A verbal presentation,
2. A written handout,
3. Flipcharts,
4. Transparencies or overheads, or
5. Slides.

Verbal presentations. Verbal presentations are probably the most informal way to share information. Points to consider about verbal presentations are:

- The speaker should be accustomed to speaking in front of groups.
- The speaker has to be able to clearly and concisely make the points that he/she wants the audience to remember.
- People have difficulty concentrating on a verbal presentation over an extended period of time, and, as a result, do not remember much, particularly specifics.

Handouts. Written handouts can detract from a verbal presentation if they are given out first. People tend to read ahead; this results in rustling papers, which distracts those who are listening. Generally, if people should follow along, give them the handout first. If people should listen and use the handout for review, give them the handout last. If a summary of the presentation will be provided, tell the audience. People are often unhappy when a summary is handed out after they have gone to the trouble of taking notes.

Figure 8.9 Male Admissions by Charge Level and Legal Status

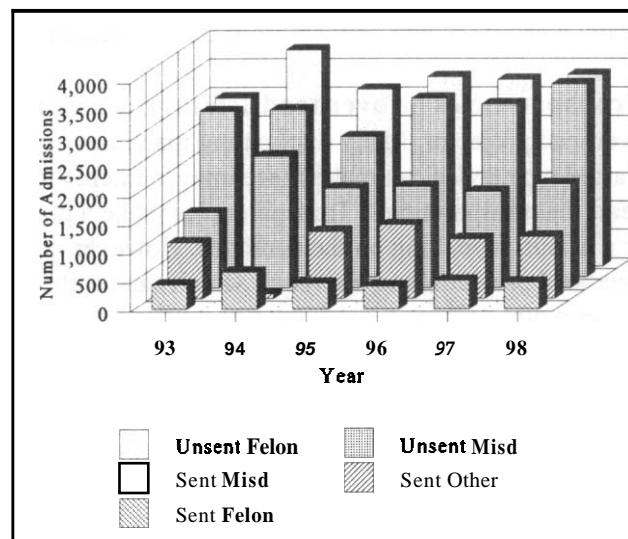
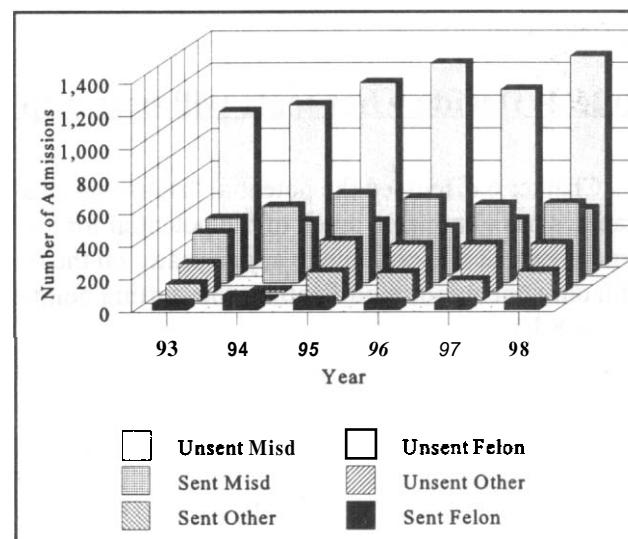


Figure 8.10 Female Admissions by Charge Level and Legal Status



Flipcharts. Flipcharts, large pads of either lined or unlined paper (usually 27" x 34"), can be prepared before a presentation or done "on the spot." Both have their advantages. Flipcharts done "on the spot" can make a presentation more informal, and the major points can be written on the chart ahead of time in pencil (light enough to be invisible to the audience but visible to the presenter). Flipcharts can introduce color into a presentation. On the other hand, they can take quite a while to prepare, and they tend to wrinkle easily, which detracts from their appearance.

Transparencies or overheads. Transparencies are clear 8%" x 11" acetate panels that are shown by an overhead projector. Many duplicating machines can copy onto the acetate from a white paper original. Varied type styles can make these images interesting, but remember that the audience must be able to **read** the transparencies. Transparencies make a presentation seem more professional, which can be a plus or a minus depending on the audience. However, using transparencies requires equipment (e.g., an overhead projector and a screen), which carries the possibility of such technical difficulties as a bulb burning out.

Slides and graphic presentations. Computer technology makes it possible to produce slides directly from graphic files and to project computer images. However, this process and the equipment can be fairly expensive. Consult the price list before you buy the software and graphic projectors because there is generally an accompanying hardware cost.

Slides are less bulky than transparencies. They require that the room be darkened, which can result in losing contact with the audience. If the presentation happens to follow a meal, the majority of the audience may doze through the presentation. Equipment failures with slides are even more disastrous than similar failures with transparencies. At least with transparencies, the presentation can be reconstructed. With the slides, even the "cheat sheets" are gone. People tend to use slides for presentations that they will do many times.

Rule 4: Decide what scale will best display the data.

Chapter 6 discussed the potential "statistical sins" associated with scale, and automation of computer graphics has eliminated much of the potential for an "unintentional" sin. However, it is important to remember that it is possible to adjust scales on the computer. Figures 8.11 and 8.12 show the same data with only one minor modification: the starting point of the Y axis is "0" in Figure 8.11 and "900" in Figure 8.12.

Figure 8.11 Admission Data on Scale #1

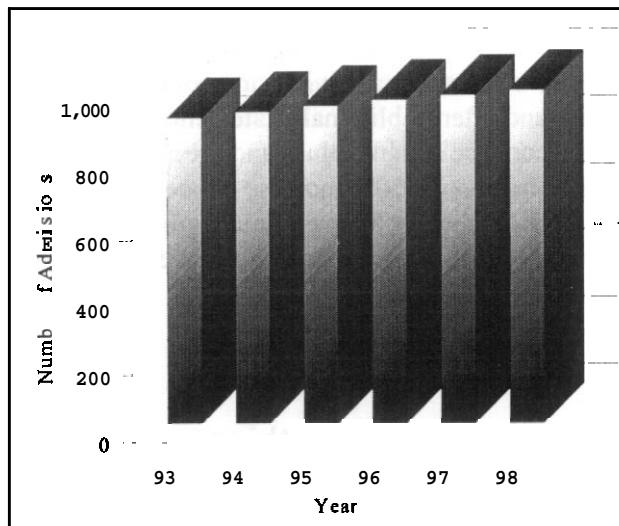
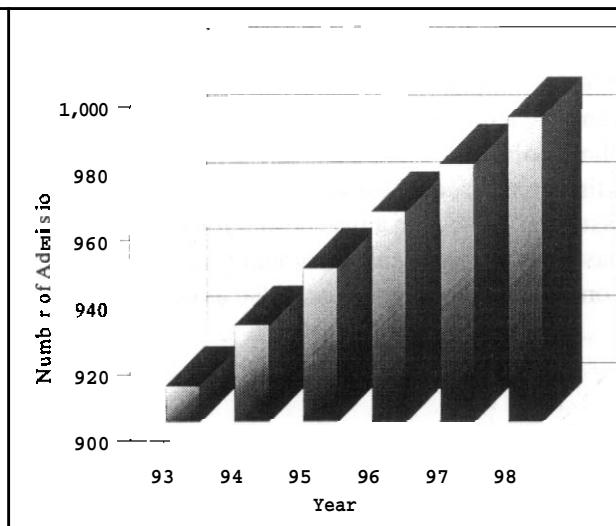


Figure 8.12 Admission Data on Scale #2



Rule 5: Consider who the audience is.

Display the data and make the presentation for the audience you hope to reach. Like statisticians and researchers, people who work in corrections and law enforcement have a language of their own that is incomprehensible to outsiders, with many "pet terms" and even more abbreviations (ROR, FTA, Part I Offenses, etc.). If the audience includes people who are not familiar with such jargon, make sure that all terms and abbreviations are explained. Show the abbreviation the first time a term is used—Length of Stay (LOS)—and thereafter use the abbreviation. It is also important to gear the presentation to the interests and abilities of the audience. A presentation for the commissioners will be markedly different from one done for the high school civics class, and both will differ from a presentation to a professional corrections association.

Rule 6: Make the graphic pleasing to the eye.

At a minimum, the tables, charts, and graphs should be neat and positioned nicely on the page. Beyond that, the following suggestions may improve their overall quality.

Draw the eye where you want it to go.

Centering something on a page does this.

Using symbols like  can have the same effect. So can using a slightly **bolder** style of type. Type comes in various sizes.

Find a way to send a message. If possible, use color to send a message. While using color in graphics can take a bit more time and money, it can send a powerful (and often subliminal) statement. Generally, "hot" colors like reds and oranges send warning messages; "cool" colors like blues and greens send calming messages. Do you want the commissioners to feel calm about a 70% increase in bookings in the last year? If so, use a pastel blue for that particular series on the bar chart. Primary colors are better than pastels; they stand out more and make a stronger point. Some colors, like pink and blue, have connotations to be avoided. Also avoid "unusual" colors, like puce or fuschia.

Rul 7: R esist he t gage. transparency. or slide.

Crowding makes a graphic image look messy and reduces its effectiveness. People process information only in limited amounts. Do not display all the data in a single table and expect the audience to read it. Make one key point per graphic. This is particularly true with transparencies and slides. Limit the number of words or numbers to a minimum.

Perhaps the best advice to keep in mind when displaying data is simply: experiment. Try different ways of displaying the data until you find a style that works for you. Sometimes, it is helpful to see what other people have done. Use the "galleries" of charts that have been created in your computer graphics program or the graphics module of your spreadsheet. Those who need to create graphics manually should see Appendix K for information about graphic tools.

Conclusion

With a little time, a little talent, a computer program, and a few dollars, most people can turn out very credible graphics. Graphics not only enhance the appearance of any report, but they also make it easier to understand the data that has been gathered. And that really is the bottom line. If people do not understand the information, they cannot use it effectively.

This manual has spent so much time and space on the technicalities of collecting, analyzing, and sharing data that a brief reminder about why we are doing this might be useful. Sheriffs and jail administrators must take strong leadership roles in using the information that is hidden in local jails to manage their organizations more effectively, plan for change, and work with other officials to improve the criminal justice system.

Sheriffs and jail administrators must take proactive roles in criminal justice affairs. For too long, jails have served as passive receptacles for the rest of the justice system. Only by using the information that often already exists can the jail become a fuller partner in the criminal justice system. This will not happen by itself—sheriffs and jail administrators have to make it happen.

In the past, sheriffs and jail administrators have been made to feel that they could not gather data themselves. This manual should have convinced you that you can. It may feel strange at first and may take time, but it can be a very powerful experience. When done collaboratively, it can help criminal justice organizations respond more effectively as a system. Information, in the right hands, is power. Systems that have gathered their own data have been surprised by how much they were able to find out about their system and were excited about how they could use the information.

Chapter 9: Getting the Most from Your Information System

What This Chapter Is—and Is Not—About

What do you think about when you hear "information system"? If you are like most people, what comes into your mind is probably "computer." This chapter is not really about computers any more than your policy on the use of restraints is about the specifications for handcuffs. However, people often want to start automating by deciding which computer to buy.

Computers are the equipment or hardware that may be used with an information system. The software used to operate the computer is the vehicle used to store and process the information. As a result, this chapter does not place a great deal of emphasis on computer hardware and software, and it will not identify products or vendors. This approach would quickly become outdated because of the rapid changes in this technology. To use computers effectively in an information system, particularly a management information system (MIS), it is important to start at the beginning—with **information**. That is one focus of this chapter.

A second focus is about asking the right questions and having appropriate and realistic expectations about your information system to make an informed decision about buying computer hardware and software. Buying the "wrong" equipment or software can quickly sabotage your efforts to automate. But when you begin the task of automating, how do you know that you are buying the "right stuff"? New hardware and software is a major investment for any jail, and sheriffs and jail administrators need to do it right the first time.

It is important to understand that computer technology changes rapidly. State-of-the-art hardware and software is outdated almost before you can figure out how to use it. It is also important to understand that the information systems currently in use in jails and their users vary greatly.

With these caveats, this chapter gives a general overview of the current technology.

A Short Course in Computers

Computers come in many shapes and sizes, with widely varying attributes. This section provides some basics about computer hardware and software.¹ In the 1990s, the amount of information about this

¹This section summarizes material that is provided in greater detail in a Bureau of Justice Assistance (BJA) monograph, *A Guide to Selecting Criminal Justice Microcomputers*, developed by the SEARCH Group, Inc., in 1990.

topic exploded to the degree that it is difficult for computer professionals to know all there is to know on the topic. Understanding some computer basics should help sheriffs and jail administrators determine the area(s) in which they need information to make an appropriate selection.

Overview of Computer Elements

All computers, regardless of size, are made up of four basic components:

- Central computer (the box), which contains the central processing unit (CPU), the computer chip that contains the memory or "brain" of the computer;
- Input devices, the equipment or peripherals used to communicate with the CPU;
- Output devices, the equipment or peripherals used to show the results of the processes that happened in the CPU; and
- Storage devices—media, such as tapes, fixed or floppy disks, etc.—used to store programs or data.

The computer may be:

- A mainframe, which is a large, central computer, typically shared by multiple user groups (such as agencies in a county);
A departmental computer, which may be as large as a mainframe or as small as a local area network (LAN), shared by multiple users within the department; or
A personal computer, which is used by only one person.

The Central Computer

Three concepts are important:

- The CPU,
- Memory, and
- Input/output interfaces.

Central processing unit. The CPU or "brain" is a silicon chip. It translates the user's instructions into a language (machine language or binary) that other parts of the computer can understand. In the course of doing so, the CPU puts data in memory, identifies where it is stored, and determines the sequence of events inside the computer.

When computer people talk about a CPU, they often use some of the following terms:

- Bits, the basic element of machine language, which is either a 1 or a 0.*
- Bytes, the number of bits needed to encode one character (e.g., an "a").
- Bus, the electronic pathway in which the CPU moves bytes.
- Megahertz (millions of cycles per second), a measurement of how fast the CPU can get data to and from memory.

In the beginning, computers were a series of switches. Each switch had only two positions, "open" or "closed." The 1 and 0 relate to whether a switch in the computer was open or closed.

These factors can make a difference in how the computer functions. In general, the more bits a computer can process at one time, the more work it can do; and the more megahertz the computer has, the faster it can go.

Memory. There are two kinds of memory:

- Random access memory (RAM) is where the computer temporarily stores data while working on it. The amount of RAM determines the total number of locations in the computer where information can be temporarily stored. Many current applications use a lot of RAM. A shortage of RAM restricts how the computer can be used. As a result, RAM expandability is a significant issue. Unless your software has an automatic back-up function, data in RAM is lost after a power failure and everything you entered since the last time you saved your file will be gone.
- Read only memory (ROM) is permanent memory and comes installed on the computer by its manufacturer. ROM stores things like the operating system.

All memory is measured in bytes. Memory is referred to in terms of kilobytes (1,000 bytes or 1 KB), megabytes (1,000,000 bytes or 1 MB), or gigabytes (1,000,000,000,000 bytes or 1 GB).

Input/output functions (I/O). If the CPU is like the brain, the I/O functions are like the body's nerves. They link the CPU with its memory, and both the CPU and memory with the computer's peripherals. Think about these devices as a switchboard.

Peripherals

If the CPU were only the "brain," then humans would be unable to communicate with it. The peripherals are the "eyes and ears" of the computer, providing input to the brain. They are also the "muscles," voluntarily or involuntarily following the brain's directions. Peripherals are divided into several categories:

- Input devices, which put data into the computer. The most common kinds are keyboards, a mouse or a pen (pointing devices), and scanners (bar code readers and optical scanners).
- Output devices, which display data in a way that either a person or another machine can read it. The most common kinds are monitors (cathode ray tube [CRT] or video display terminal [VDT]) and printers.
- Storage devices, which store data (remember that RAM is temporary). The most common kinds are disks (floppy diskettes, hard disks, and compact disk read only memory [CD-ROM] devices) and magnetic tape. The most significant issue is how **much** of this is needed.

Communication Devices

Communication devices are the mechanisms that let the computer communicate with the **peripherals** and other computer systems. There are four types:

- **Input/output ports**, the interfaces used to connect the wires between the central computer, the peripherals, and other communication devices. Ports can be either serial or parallel, depending on the number of bits that can be sent between a computer and peripheral at one time. Parallel ports, usually used for short distances, can send multiple bits at one time using a wire for each bit. Serial ports take the bytes apart, send them one at a time in a stream, and then put them back

together at the other end. Serial ports require some type of checking process to be sure that the bits are reassembled correctly. They are used over long distances.

- Modems allow the computer to send data over a telephone line. They are measured by the rate at which they can send data (baud)—faster is better. Modems require special communications software.
- Facsimile cards or fax modems allow the computer to function as a fax machine.
- Networks or local area networks (LAN's) allow a number of independent devices to communicate with each other. This, in turn, allows computer users to share software, peripherals like printers, and data. In a LAN, the computer on which the shared data is stored is called the "sewer."

Overview of Software

If the hardware described above is the equivalent of the body's nervous system, then software is thought itself. It consists of electronic or magnetic signals that tell the elements of the computer what to do. Software can be divided into two basic types:

- System software, including the computer's operating system (such as DOS or Windows), and utilities, tools that support the operating system. Utilities help a user perform routine file management functions.
- Applications software, which are the instructions that allow the user to work on the computer. There are many applications, including word processing, spreadsheets, and data bases.

Applications Software

Software may be proprietary (owned by a corporation or individual and protected by copyright), public domain (not owned by anyone and available at no charge, for example through the Internet), or shared (copyrighted but able to be shared, although often without printed documentation). Of the three, proprietary software is the most significant.

Proprietary software comes in two basic types:

Software, such as word processing, spreadsheet, data base, and communications packages, designed to be used by different disciplines; and
Software designed to be used for a specific function, such as fleet management, commissary, or jail booking.

Data base applications. Data base software is the most commonly used type for information systems. It stores and retrieves information in a structure and provides a way to examine relationships among the data elements. It is important to understand that there are different kinds of data bases:

- File management/report generator/flat file data bases allow the development of various reports from each case in the file,
- Relational data bases allow the development of reports from multiple files, and
- Hierarchical data bases store the data in several levels rather than in a "flat" file.

File management data bases, like address files, are easy to create and use, but cannot be linked to other data bases or files and therefore support limited reporting. Relational and hierarchical data bases combine and compare data from various records. They are complicated to create, but provide much more powerful and flexible reports.

Data base software organizes data elements into "fields," in which related information about a single case (or "record") is stored. The fields can be summarized in standard or specialized reports. Fields are organized into data base files that logically belong together, such as booking, classification, housing, commissary, etc. Data bases routinely allow users to:

- Browse—scan records in the data base for information,
- Search or query—find records based on specific criteria (such as pre-trial felons who were in custody for more than 6 months),
- Sort—list records in a specified order (such as an alphabetical listing of inmates who had more than five disciplinary violations in 1995), and
- Update—add, change, or delete records in the data base (usually a function restricted to a limited list of specified users).

At first, data bases were not particularly "user friendly." However, the "kinder and gentler" data bases of the current software generation can:

- Easily develop the contingency tables discussed in this document;
 - Allow a moderately experienced user to develop special reports, in addition to the standard reports, in response to the need for additional information;
 - Perform basic math operations, such as totals, counts, minimums and maximums, and standard deviations;
- Create graphics;
Import and export data to and from other applications, such as spreadsheets and some word processing programs.

A "user friendly" off-the-shelf data base software package is preferred for specialized data collections such as the Inmate Profile Data Collection

Spreadsheet applications. If data base software is the "file cabinet" of the computer world, spreadsheet software is the ledger. Spreadsheets are useful in organizing data that requires statistical analysis or a lot of computation. Spreadsheets were originally developed as a way of dealing with accounting information, but now most also have some data base qualities, although they are limited, for example in the number of criteria you can use to sort the file and identify particular cases. Most spreadsheets have a graphics component that allows the development of bar charts, pie charts, and the other graphics discussed earlier. Most newer versions of spreadsheets have the ability to import and export (or "translate") spreadsheet data into other formats, such as those that could be read by a data base or word processing program. A "user friendly" off-the-shelf spreadsheet package is the preferred option for Inmate Population Data Collections.

Communications software. Communications software is not typically used to collect or analyze data, and yet it is some of the most common software found in criminal justice systems' computers. The way in which agencies communicate with the National Crime Information Center is a form of communications

software. Communications software is used to exchange data electronically on an electronic bulletin board or via the Internet.

Word processing software. While word processing software is probably the most common application found in most organizations, it has little use in collecting and analyzing data, except in developing the report. It should not be used in a major data collection or as a recordkeeper, even with the math functions that some word processing packages provide. It is very useful, however, to have a word processing program that can link to a data base or spreadsheet through its table function or create graphics through an internal drawing or chart-making program.

A last word on using applications. Particularly at the management level, most people find they use a combination of software packages for different aspects of their information systems, data collection, analysis, and reporting functions. As software has become more integrated and more flexible in its ability to move data from one type of application to another, that process has become easier. If your information system and other software does not allow this, it will likely impede your analytical efforts. These limitations may be an indication that it is time to upgrade your applications software.

A Last Word on Hardware and Software

After this discussion, it should be obvious that many choices must be made in selecting hardware and software. How do you know which hardware and software to buy? The answer is both very simple and very complex. It depends on what information you need, and answering that question thoroughly and effectively will make the difference between a system that supports your operation and one that handicaps it.

Getting the Most Out of a Data Base

If you already have a jail information system that was designed for your operation, regardless of whether it operates on a mainframe, a mid-size computer dedicated to your department, a server, or a personal computer, it is probably a data base of some type. If you do not know what it is and what it can do, make the effort to find out!

A good way to begin is by finding the Data Dictionary for your data base. It is similar to the code books that were described in Chapter 4. A Data Dictionary is:

- An organized reference to the content of a data base;
- A tool that helps to manage data and convert it into information; and
- A tool that defines the characteristics of each data element in the data base by providing the name of the data element, its definition or aliases (or other similar terms), a physical description of the data element (i.e., its length, type, format, and a rationale for any rules for entry, etc.), the edit rules (acceptable range or reference values for the data element), its function (a description of how this data can be used), responsibility for the data element (i.e., who creates and maintains this data element), and access authority (who has the authority to have access to this data element).

If you do not have a Data Dictionary, consider developing one as a way of first learning how the system works and then improving its usefulness. Several guidelines can make this process (and the resulting data entries in the information system) more useful.

- Avoid ambiguities. You may recall the discussion in Chapter 2 of the variable Charge and the problems associated with defining it. "Capacity" is another term like "charge." It may mean a court-ordered capacity, a rated capacity (established by the state), a design capacity (the number of beds provided in the design), actual capacity (the number of beds normally available), or working capacity (actual beds less any temporarily unavailable beds). Be specific.
Document aliases. A number of terms can refer to the same thing. In a direct-supervision jail, we might call the areas in which inmates live modules, housing units, or pods. Identify which terms mean the same thing and decide on the preferred term.
- Use naming conventions. Name your data elements consistently (e.g., BookNum and FBINum, not Booking# and FBINo).

What to Do Before Buying or Upgrading a Computer

Sheriffs and jail administrators who want information systems to help in decisionmaking, not just as an automated recordkeeping system, have a number of tasks to do before collecting vendor information. These tasks are described in detail in the BJA monograph, *A Guide to Selecting Criminal Justice Microcomputers*, and were reviewed in Chapter 2 of this manual.

Before buying the computer and software, your information system needs should be defined in more detail. There are three steps in the needs assessment process:

1. Define the problem,
2. Review the existing systems or functions, and
3. Document the relationship of this system and information.

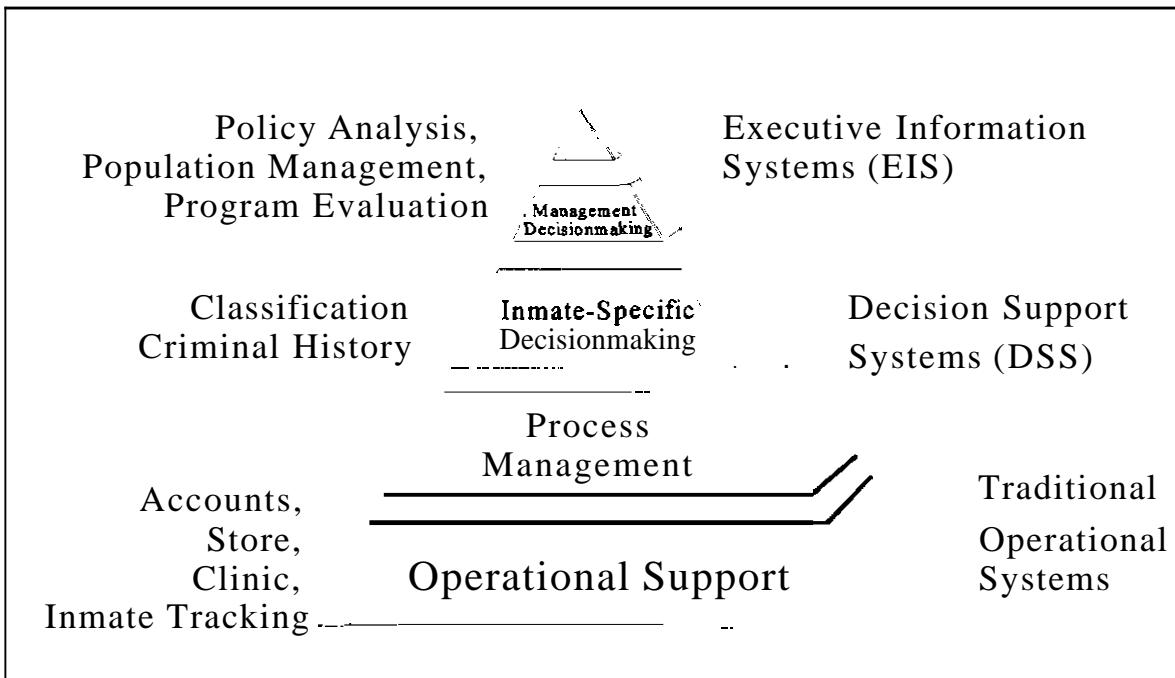
Determine the Purpose of the Information System

One difficulty in discussing jail information systems is that they mean different things to different people. Figure 9.1 illustrates three types of information systems (Executive Information, Decision Support, and Traditional Operational), the types of activities they support, and examples of functions typically automated in each.

Many jail information systems currently fall into the category of Traditional Operational Systems. They record and store information about bookings, inmate accounts, commissary transactions, etc. This manual suggests that sheriffs and jail administrators need an executive information system (EIS) to provide timely, accurate, and relevant information to determine the status of various activities and assist them in both long- and short-term planning. A good EIS can:

- Determine the status of individual programs,
Monitor organizational performance,
- Analyze the cost of delivering various programs and services, and
- Measure the impact of legislative and policy changes.

Figure 9.1 Jail Information Systems



The EIS grew in popularity in the 1980's as the correctional environment became more complex and computer technology became more accessible. Typically, early EIS's were tied to a mainframe by expensive, specially written computer software. These systems took information from an already existing data base and made it accessible to managers, often in summary form. However, with better, quicker, and more powerful personal computers (PC's) and LAN's, EIS has come into the mainstream. The tools are currently available to build an EIS using a good, contemporary data base program and other software tools. Once the statistical data needed is defined, the basic reports can be set up in an EIS. The capabilities of the data base software selected should allow a query of the data base to find answers to specific questions not covered in the routine reports. The data base also can be structured to alert you to information that requires your attention, such as any housing unit in your facility that is over capacity.

Define Your Needs

If these functions are wanted or needed, a system will need to be capable of moving beyond the traditional jail information system to an EIS. Begin with a good definition of the problem.

Step 1: Define **the** problem. Chapter 2 discussed defining a problem. That is where the needs assessment begins. At this point, it is productive to involve people who have hands-on experience with the problem. They can bring to the problem statement the "nitty-gritty" details that administrators may not know. Developing the problem statement helps to focus the assessment on specific areas or functions. Consider the following example.

The administrator of a 100-bed jail learned that Frank E. Felon, an inmate back from the state system, had filed yet another lawsuit, alleging that he had been denied recreation for more than a month. The jail

administrator knew that there had not been a disciplinary action against the inmate that resulted in the loss of recreation. The jail could prove that Felon's unit had been scheduled for and received recreation during the month in question and, from the shift log, that a number of inmates had gone to recreation. However, no one could identify which inmates, nor could the recreation officer remember if Frank Felon had been to the gym. The department, then, could not prove that the allegations were untrue. The jail administrator, shift supervisor, and recreation officer met to discuss the situation.

The existing system for logging inmate movement does not identify inmates who are moved to activities like recreation, and the current system of manually listing inmates is too labor intensive. As the recreation officials said, "I can either do recreation or write that I've done it." But the jail administrator wanted to avoid any further trips to federal district court, so something had to be done.

Step 2: Examine the existing system. Examining the existing system (defined here as the Program Participation System) could be a very poorly defined task. But jails have an "edge" on a number of other kinds of organizations because they have policies and procedures. Reviewing these should provide both the purpose of the system and the "flow" (the step-by-step listing of how the procedure is carried out). By using a computer programming technique called systems analysis, you can relate the steps in the procedure to informational needs.

Systems analysis looks at all processes in terms of inputs, throughput, and output. In trying to track information, it is important to think about activities in terms of the information they generate, modify, or capture. Inputs start the system. In this case, the inputs are the recreation schedule and an inmate's request to go to recreation. Outputs result from the steps in the process. In this case, the outputs are entries in the daily log that a number of inmates went to recreation and documentation of the inmate's participation in the recreation program. Throughput is the steps or processes in the program participation process. In this case, they are:

1. Log the inmate out of the housing unit,
2. Log the inmate into the program area,
3. Program participation,
4. Log the inmate out of the program area,
5. Log the inmate back into housing,
6. Summarize the hours of recreation provided.

Figure 9.2 Inmate Programs Flowchart

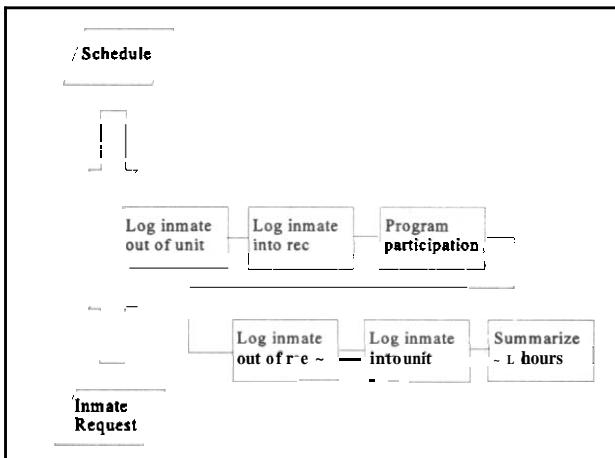
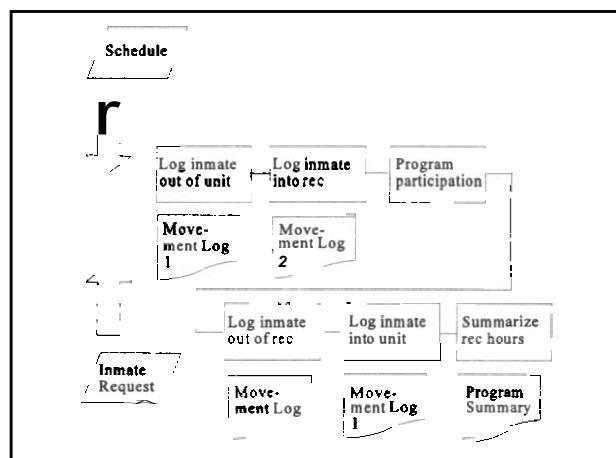


Figure 9.3 Inmate Programs Flowchart with Documents



Systems analysts use a graphic shorthand called flowcharting, shown in Figure 9.2. Each of the rectangles represents a step in the process. The parallelograms represent inputs. Figure 9.3 shows how to relate this to information gathering by showing the documents that are generated as a result of the process.

Once the flow is described, several other things should be considered:

How many people use or modify the information recorded on the documents?

- What is the volume of recorded information?
- How is the information processed?

In this system, three basic documents store information about program participation:

- Inmate Movement Log 1, completed at the housing unit;
- Inmate Movement Log 2, completed in the program area; and
- Program Summary, completed by the person responsible for each program

As the jail administrator and staff continue to examine the issue, additional issues surface:

- It takes too long to manually record the name of each inmate who leaves or enters the housing unit in Movement Log 1 and each inmate who enters or leaves the program area in Movement Log 2. There are too many inmates, and they move too often.
- The recreation officer has to manually calculate the number of hours of recreation delivered each month.
- The existing forms still cannot meet the jail administrator's need to prove exactly who participated in what program.

Step 3: Prepare a needs assessment statement. The jail administrator can now prepare a needs assessment statement that provides more information to make a better choice about how automation could be used to solve this problem and what approach might be most effective. For example, it is clear that:

Just putting the existing inmate movement logs on the computer will not solve the problem if the housing and recreation officers still have to type the inmates' names. As a result, the system must minimize manual entry of inmate names through the use of such technology as **barcoded** inmate ID bracelets or ID cards.

- Unless Movement Log 2 is linked to the Program Summary, the recreation officer will still be forced to manually calculate the hours the program was delivered.
- Unless the software used for the movement logs can allow the jail administrator to **query** the system to identify each time Frank E. Felon went to recreation, all the movement logs will have to be reviewed for the period in question.

Unless the system can keep the information for an unspecified time period, all of the documentation about inmate movement will not help in court.

While this is just the beginning of the process to follow to determine what kind of system best meets your needs, it should be clear that much of the work involves understanding the flow and volume of information in the facility.

Step 4: Determine the data elements. Chapter 4 suggested a review of forms and logs be done to see what information is collected where. For our Program Participation System, one report and two logs provided the information. The next step in the process involves looking at each of these forms to identify:

- The data elements that are currently included,
- The data elements that are **not** currently included but should be, and
- The information that is included in both the report and the log.

Step 5: Estimate the volume and storage requirements. Once all the data elements have been identified, you will need to estimate the volume of your information—how much data must be handled and the speed at which the computer must operate. In the case of the movement log, it would be important to **know** how many inmates move at one time and how often during a shift, day, or other time period movement occurs.

One of the most significant aspects of the information system is its storage devices. A way to begin to understand storage requirements is to go to the paper files where the information is stored to determine how much paper is generated over a time period. While recognizing that it is really the number of characters on each piece of paper that will help with storage size, a quick look in the files should give you an idea of how much will need to be stored. Related to storage is the question of whether you will need access to this information immediately or if it can be stored on external storage devices, such as floppy disks or tape.

Step 6: Determine the reporting requirements. The next step is to determine what the system should tell you both routinely **and** on an "as-needed" basis. For example, a standard report this jail administrator would want is the Program Summary, which should be designed to identify how many hours of recreation were provided, the number of inmates who went to that service, and perhaps what proportion of inmates do not use the service. On an "as-needed" basis, the jail administrator would want to be able to generate a report that listed all inmates who went to recreation during a specific period.

This is a very difficult task, because information needs change over time. As a result, systems that are very flexible, that allow links to information from multiple tables or files, and that put together information in varied ways on an "ad hoc" basis are likely to work best for management decisionmaking.

Step 7: Explore compatibility issues. Avoid creating "information islands," which are isolated and cannot communicate with each other. The jail administrator will need to consider whether the system to track inmate movement to recreation should interface with any others and whether these issues are applicable to any other systems. A review of other program functions and logs indicated that the problems associated with movement logging were much broader than just recreation. Any solution created for recreation movement would apply to any other programs and also to transfers within the facility and cell assignments. Because cell assignments and transfers were recorded on the daily roster, currently created by the jail's booking system, the issue of compatibility with the jail's existing software was raised. And by broadening the purpose of the system, it was possible to avoid a very common automation phenomenon—an underestimation of its use or applicability.

Step 8: Explore computer options. This is where most organizations start in the process of developing an automated information system, and it is where this section of the chapter will stop. For additional information about automated systems, contact the NIC Information Center (1-800-877-1461).

System Expectations and Guidelines

When the first edition of this manual was published in 1982, automation was uncommon in jails. Computers seemed mysterious and the expectations of automation were high. In the late 1990s, the computer has lost its mystique, and many law enforcement and corrections agencies routinely use computers for office and communication functions. However, relatively few agencies have completely integrated computers into their daily operations to capitalize on the vast information resources in their systems. If there is any blame to attribute in this situation, it is shared. Many automated systems are not "user friendly," and many agencies continue in the unfortunate belief that computer work is clerical. This section offers some guidelines for automated systems and their users.

The User Manages the System

A good system does not require programmers to reprogram the computer to modify a report or to obtain anything other than standard reports. In older "unfriendly" systems, the user is so dependent on the programmer that he/she becomes a receiver, not an analyzer, of information. Particularly in network environments, users need an "in-house" network manager, who, as a "high-level" user, can generate and develop special reports, make minor modifications to reports, and in general use the system.

The System Avoids Duplication of Effort

A good system can help to avoid duplication of effort. First, once a variable, such as Inmate Name, is entered, it does not need to be reentered on subsequent reports. For example, once the initial booking screen is completed, the inmate name and other pertinent identifying information should automatically appear in later forms, like health screening. Second, the information is entered directly into the computer, not written on paper and subsequently entered into the computer. Think about it! By setting up a separate anti parallel paper system, we increase the level of effort.

The System Is Reliable

System reliability has a direct relationship to the tendency of many organizations to maintain a parallel paper system. If the system is often down and requires someone from data processing or another agency to reset it, people begin to mistrust it. After all, jails never sleep--and neither should the computer. Reliability in large computer systems can be absolute; "double fault tolerant" systems, which never go down for service and which have multiple redundant parts, exist. In small computer systems, reliability means good and regular back-up procedures.

The System Has a Safety Net

Many new users fear that they will somehow "lose or modify" all the data. And if the automated system is your records system, you need to be concerned about the integrity of the records. A "safety net" can take many forms, such as:

- Automatic back-up features, which preserve data in its previous form;
- Data that is modified or replaced never replaces or overwrites the initial entry;
The operator works with or manipulates a copy of the data, not the original.

It is critical to understand the "safety net(s)" your system provides.

Match the Right Application to the Task

As the section on software in this chapter suggests, each kind of software has a specific purpose. However, people often use the software that is available, regardless of whether it is appropriate for the task. This can lead to user frustration. For example, one jurisdiction entered inmate count statistics into a word processing document and then was unable to perform any math functions. Another jurisdiction entered a tool and toxic materials inventory into a word processing document that could not be sorted as needed.

Know How the System Handles Updated Information

Many automated systems do not handle changes to original information well. In many cases, the original information is written over and lost. This can present a number of operational problems. For example, one of the things that often changes is Bond Amount. At the initial appearance, the amount is set, but bond is commonly lowered at some point in the process. Imagine trying to determine whether the level at which **bond** amounts are initially set is unrealistically high. Without knowing that the amount in the computer record is a revised amount, things might look normal. However, the user can require that the computer store both initial and revised amounts.

Format Flexibility

As managers, most sheriffs and jail administrators want summary information on a regular basis with the ability to view the detail if needed. Some rigid information systems have to churn out all of the detail to generate a summary. In one jurisdiction, this resulted in printing a 400-page report daily to get to the 1-page summary the jail administrator needed. In addition to the waste of paper and resources, this process took approximately one hour of computer time, a significant issue on a county mainframe.

Determine Internal Checks

It is easy to make entry errors. Good computer software has one or more of the following characteristics:

- It minimizes the amount of direct entry by using things like "pop-up" menus for the most common answers.
- It asks the operator to **confirm** an entry.
- If the operator enters a numerical code, it defines in words what was entered.
- It allows only numerical characters in numerical fields.
- It designates critical variables as "forced" fields (which must have an entry before the program will continue) to ensure that they are entered.

Such internal checks are necessary to ensure that the information is entered correctly.

User Guidelines

Regardless of the type of hardware and software used, the most important component in the automated information system is the user. Without a creative and curious user, the best system will never live up to its capabilities. As a result, the last word in this manual is about people, not machines.

Explore the System

One of the best ways to learn about your system is to see what it can do. To do this safely, make sure that you have a safety net to protect the integrity of your data before you start. Try to generate special reports. Try to analyze information on your computer. Try different applications before you start. Try to create reports on the computer that you now do manually. Explore each of the functions and applications available on your computer. If you do not try these things, the computer will continue to generate only the standard reports.

Practice Analysis

In any information system, just as in statistics, the user's responsibility is analysis. It is up to users to:

- Determine what the information implies,
- Search out relationships among pieces of information,
- Identify missing information, and
- Use the available information to manage and direct the organization

Analysis gives us the tools to be detectives. In information systems and statistics, we may use different tools, but the thought processes are similar. Too often, our feelings about computers or statistics cause us to overlook these resources that are rich in potential clues to organizational problems.

Appendix A: A Glossary of Statistical Terms for Non-Statisticians

Aggregation is a process of grouping data together. The reverse is called disaggregation. Experimenting with different ways to group data is one way that analysis takes place.

Analysis of Variance is an advanced statistical technique used to determine whether significant differences exist between three or more groups.

Audit is a form of process evaluation to determine whether resources are being used for the intended purpose **and/or** the intended target population. It may include the elements of process evaluation, program monitoring, and effectiveness enumeration. It includes internal controls designed to ensure that program goals and objectives are defined and achieved and that the program is in compliance with statutes and regulations and is economically effective.

Axes are the vertical and horizontal lines that orient many charts and graphs. Statisticians call each line an **axis**. The horizontal line is the **X axis**, and the vertical line is the **Y axis**.

Barriers are conditions or obstacles that must be overcome to solve problems.

Bias is built-in error. It usually comes from picking a bad sample or from the perceptual bias of the person designing the study.

Case can be a person, event, or object about which **information** is being gathered.

Case Identification Number is a unique number assigned to each case.

Cell is the name given to each "box" in a contingency table.

Chi Square is a statistical test that determines whether being a member of a particular category of one data element has anything to do with being a member of a particular category of another.

Classes are the groups into which the data is divided to calculate and display frequency distributions.

Cluster Sample is a sample in which the population is clustered into specific areas or units (e.g., housing units). Then the units to be sampled are selected randomly, and information is collected about each case in the unit.

Code Books translate the data elements and their categories (which are usually written words) into numbers so that computers can process them more quickly. Code books identify the data elements, the

categories for each data element, the number of digits in the code for each data element, and (if the data collection is computerized) the column **name/number** or field name of each data element.

Cohort Sample is essentially the same as a cluster sample, but is based on time periods rather than physical or geographic areas (e.g., all prisoners arrested on a specific, randomly selected date).

Columns are the vertical lines in statistical tables.

Confidence Intervals are statistical limits. In sampling, statisticians use probability theory to determine what the odds are of developing a statistically similar sample. Most statisticians want to be at least 95% certain that they would get similar results if they analyzed another sample, but some want to be 99% certain. In other words, they want either 95% or 99% confidence intervals around their samples.

Contingency Table is a chart that displays the relationship between two frequency distributions. They are used in calculating the Chi Square statistic and are often called cross-tabulations.

Correlation measures the degree and direction of the relationship between two data elements. Correlations are positive if change in one direction in one variable is associated with change in the same direction in the other variable. These variables have a direct relationship. Correlations are negative if change in one direction in one variable is associated with change in the other direction in the other variable. These variables have an inverse relationship.

Cross-Sectional Data Collections are done only once; most typically they provide information about one time frame. They can only describe something.

Curvilinear Relationships are not uncovered by tests of correlation. A curvilinear relationship is found in a situation like the following example: When taking tests, people who have either very low or very high anxiety levels tend to get low scores, but people who have moderate levels of anxiety get high scores.

Data is information.

Data Collection Sheets are forms on which people record the information being collected. Statisticians call them **Instruments** or **Survey Instruments**.

Data Element is a piece of information. Statisticians and scientists call data elements "variables."

Data Points or **Scores** are the individual measurements for a single data element or variable,

Decision Tree is a series of questions that funnel to the only possible answer or conclusion.

Degrees of Freedom is a mathematical concept that measures the amount of information available in normally distributed data.

Dependent Variables are data elements that researchers are trying to explain (indicated by "X") by analyzing their relationship to a known variable.

Descriptive Statistics summarize information about a population or a sample

Design is the strategy of an evaluation, specifying those individuals or units to be studied (sampling procedure to ensure representativeness), the measures to be used in the collection of data, and the comparison groups to be contrasted to determine the effect of an intervention and the outcome needed to indicate program impact.

Empirical Rule is a statistical law or theorem that specifies that, if the data is distributed in a normal or bell-shaped curve, 68% of all cases will be found in the interval created by adding and subtracting one standard deviation to the mean. If two standard deviations are added to and subtracted from the mean, 95% of all cases will be found. If three standard deviations are added to and subtracted from the mean, then all (or nearly all) of the cases will be contained in that interval.

Evaluation is the process used to address the effectiveness of a program. An evaluation can be of the process used to implement the program or of its outcome.

External Validity is a research problem stemming from having information about **part** of a population. The question that arises is whether the information discovered in the research about **part** of the population is equally applicable to the **whole** population. This is often phrased in terms of the relative ability (or inability) to generalize the results to the whole population.

Frequency Distribution is a statistic in which the responses to a data element are categorized and the number of responses in each category are expressed in percentages.

Goals are end results to be achieved.

Hypothesis is an educated guess about what may be involved with or cause the problem. It is often expressed as a question and is what researchers try to prove (or disprove). Researchers complicate this by a series of complex logical moves that allow acceptance or rejection of something called the "null hypothesis," which is the opposite of the statement just defined. For our purposes, just think of the hypothesis as an educated guess.

Impact usually refers to the difference in outcome between an intervention and no intervention. It can also refer to the contrasting differences between several alternative interventions.

Independent Variables are variables that researchers think may cause or have some sort of relationship with "X (the dependent variable). Independent variables are indicated by "Y."

Inference, in statistics, means that what has been discovered about the sample is assumed to be true for the population. If the sample was collected randomly and is large enough to adequately represent the population, making an inference is acceptable statistical practice if the reliability of the inference is identified.

Information Systems are structurally or functionally related mechanisms and procedures that provide information to an organization, usually in summary form. They may be automated (computerized) or manual (on paper).

Internal Validity determines whether there is any proof that what was done to change a situation actually caused the change.

Intervening Variables act as a link between or interfere with the interaction of dependent and independent variables.

Level of Measurement refers to how different types of data elements can be measured. Data elements are divided into three basic categories. The first group (whose answers can be divided into a series of categories) is called nominal data elements. The second group (whose answers can be ranked in some kind of order) is called ordinal data elements. The third group (whose answers are expressed in equal units) is called interval and ratio data elements; this group is further divided into two subgroups, depending on whether zero is a possible response. The level of measurement of each data element determines the types of statistics that may be used to analyze it.

Linear Relationships usually occur when a strong correlation exists between two variables. If a graph is plotted to show all the data points, a strong positive or negative correlation will result in data that "clusters" together. A straight line can be drawn through the cluster to summarize how the data points lie in relationship to each other. When it is easy to draw such a straight line, the relationship between the data elements is said to be "linear." This line is called the trend line and is used by statisticians for a special kind of correlation called population forecasting or trend analysis. In population forecasting, statisticians attempt to determine whether a relationship exists between a data element, such as a jail's average daily population, and time.

Log Book is a chronological list of events or activities

Longitudinal Data Collections are those in which the same data is collected from year to year, allowing comparisons, documentation of change, and (possibly) identification of reasons for change.

Marginals are the numbers written around the outside of a contingency table.

Matched Groups exist if each case in Group I is paired with a case in Group II with common characteristics. The most likely instance of matched groups in correctional research is when one group is compared **before** and **after** participation in a program.

Mean is a measure of central tendency in which all the measurements for a particular data element are added and the result divided by the total number of measurements in the set (the "mathematical average").

Median is the middle measurement when all of the measurements in the set are arranged according to size. If there are an odd number of measurements in the set, the median falls exactly on the middle measurement. If there are an even number in the set, the median falls exactly halfway between the middle two measurements.

Methodology is the method used to implement a design. Methods usually can be divided into several categories: descriptive; process and outcome or naturalistic observation; correlation; quasi-experimental; or experimental. No one methodology is correct, but the choice of methodology depends on the nature of the research question and the time and resources available for the research project.

Mode is the measurement that occurs most frequently

Model is a suggested approach to address a problem that is considered appropriate or ideal.

Normal or Bell-shaped Curve is the most common or usual probability distribution. If the total area under the curve is set equal to 1, then the area within one standard deviation of the mean is approximately 0.68; within two standard deviations, 0.95. Normal probability distributions are symmetrical around the mean. The location and spread of the normal distribution will depend on the values of the mean and standard deviation.

Objectives are concrete items that can be measured to determine if a goal is achieved. Researchers often link a series of objectives into an **Operational Definition** of a goal or concept.

Population, in a statistical sense, means all the units (people, items, events, etc.) from which a sample can be drawn.

Population Data Elements are variables for which information is kept on each person, event, or object being studied (the whole population). These are elements that should be kept routinely.

Pre-Test is a process in which a **Survey Instrument** is given to a small group, similar to the targeted population, before the main data collection to determine if there are problems with it.

Probability allows inferences about a whole population to be made from a sample.

Probability Laws (or **Probability Theorem**) allow analysts to calculate the chances of drawing a particular sample.

Problem Statement is a description of the situation clear enough to allow analysts to identify the pieces of information needed to determine causes and suggest solutions.

Proportional Random Sample divides a population into groups, and cases are selected randomly from the population until the groups are in the same proportion in the sample as they are in the population. If a population is composed of 40% felons and 60% misdemeanants, the sample would be also.

Qualitative Research Methods (most commonly interviews and observation) describe an event or phenomenon in words.

Quantitative Research Methods rely on numbers (most commonly statistics) to reveal something about an event or phenomenon.

Random means that each case has an equal chance of being selected for a sample.

Random Number Tables are charts of numbers without any logical sequence.

Rank Ordering is a procedure to arrange data elements from lowest to highest

Reliability is a statistical concept that addresses how precise the sample results must be or whether the same results could be achieved if the measurements were done again.

Research is a systematic method of addressing an empirical question. Experimental research is considered the most sophisticated method.

Rows are the horizontal lines in statistical tables

Sample is a specially selected part of a population

Sampling Schemes (or **Sampling Frames** or **Sampling Designs**) are ways to determine which cases go into the sample.

Scales for graphs are like the scales on maps, where one inch represents so many miles. In most cases, the distance on the scale represents a fixed number of cases.

Scores or **Data Points** are the individual measurements for a single data element or **variable**.

Simple Random Sample is one in which cases are selected randomly from a homogeneous population (usually by using a random number table or similar tool).

Skewed Distribution is a normal or bell-shaped distribution that "leans" more toward one end of the scale than the other.

Solutions are ways to fix a problem. Every problem has more than one solution. If there is **no** solution or only one solution, the result is a constraint that limits how operations or choices can occur.

Spurious Correlations are caused by chance. For example, a very strong positive correlation may be found between a county's jail population and the price of hog futures on a commodity exchange, but it is unlikely that the two situations are related.

Square is a mathematical function in which a number is multiplied by itself.

Standard Deviation is the square root of the variance. It measures the degree of variation found in the data. See also **Empirical Rule**.

Standard Error of the Mean measures the similarity of the samples (how much variability exists in the data).

Statistical Sampling allows information to be gathered from a segment of the entire population in such a way that the information can be applied to the entire population.

Statistical Significance describes whether a difference between the measures of a single data element in two groups is likely to result from sampling differences or if it is a true difference. If a score is outside two standard deviations in either direction from the mean, it is "statistically significant at the .05 level." To be "significant at the .01 level," the interval has to be plus or minus 2.68 standard deviations from the mean.

Statistics, as a field, is an area of mathematics with strong ties to logic. It uses numbers to make large or diverse amounts of information more understandable. Statistics are used for four main purposes:

- To summarize information,
- To determine how seriously to take differences between groups of people or events,
- To determine how strongly pieces of information are related to each other, and
- To predict future trends.

Stratified Random Sample is one in which the population is divided into groups, such as felons and misdemeanants, and equal numbers of cases are selected randomly from within each group.

Survey Instruments are the forms or formats in which people record the information being collected.

Symptoms are visible indications that indicate a problem is present.

Systematic Sample is one in which information is collected about cases in a certain sequence (e.g., every third person booked) until the number of cases required for the sample is reached.

Tally Sheet is a form listing all the categories for each data element. A check mark is placed next to the appropriate category for each case.

Tests of Significance are statistics that examine the differences between two or more groups

Variability is a statistical concept that determines how different the individual cases that comprise an entire population are from one another.

Appendix B: Annotated Bibliography

This bibliography is divided into two sections:

- Data collection or analysis documents developed specifically for the criminal justice system in general and jails in particular and
- Documents found to be useful both in preparing this manual and in learning about research methods and statistics. Also identified are documents that will be particularly helpful to people who are not planners, researchers, or statisticians.

Many of the criminal justice documents can be obtained from the NIC Information Center, 1860 Industrial Circle, Suite A, Longmont, Colorado 80501, telephone 1-800-877-1461. Citations for resources not available there include the name of publishers, which local bookstores may need to order them.

Some of the documents are older, but statistics and research have not changed much since this manual was originally published in 1982. However, because information systems technology has changed greatly and rapidly, most of those resources are dated. As a result, the resources listed here focus exclusively on data collection and analysis, statistics, and research methods.

Resources for the Criminal Justice System

Assessing the Effectiveness of Criminal Justice Programs. Robert A. Kirchner, Roger K. Przybylski, Ruth A. Cardella. Published by the Justice Research and Statistics Association and the U.S. Bureau of Justice Assistance in Washington, D.C., 1994.

This handbook aids criminal justice policymakers and program managers in assessing the effectiveness of their programs. It identifies general criteria for program effectiveness and develops questions to be answered during an evaluation.

Corrections Planning Handbooks, Book 3 (Collect Data). Prepared by Farbstein and Williams Associates for the California State Board of Corrections, 1981.

California developed a needs assessment process that its counties had to follow when applying for funding for local jail construction. One step in the needs assessment process was to collect data about local jail operations. Book 3, developed to assist counties with data collection, is an excellent resource for counties that will undertake a major inmate profile data collection. It gives detailed, step-by-step instructions about how to profile the jail population, existing programs, and current criminal justice system functioning; how to document trends in the justice system and jail population; and how to convert these projections to capacity and program needs. This approach provides for the integration of alternatives to incarceration in the planning process and provides excellent guidance in assessing criminal justice system functioning. This document is written in

understandable terms and is geared toward counties that do not have access to computers for processing data. Its one possible drawback is that it is based on a given set of data elements. Systems that want to collect other items may experience a little difficulty integrating them into the study. A word of caution about this manual. If you are going to use this approach, follow it step by step. Do not **skip** around. All in all, this is probably the most comprehensive resource that has been written for jails. It is based on a solid planning methodology, and its potential for use in other States is high.

Criminal Justice Analysis: An Introductory Course. Henry S. Dogin, Perry A. Rivkind, and Richard N. Ulrich. Prepared for a special seminar of the Law Enforcement Assistance Administration in 1978.

This document consists of fact sheets and exercises put together to train seminar participants in using statistical information for criminal justice policymaking and planning. It will not tell you **how** to analyze the information, however, or give you the answers to the exercises that are included.

Guide to Data Collection and Analysis: Jail Overcrowding/Pretrial Detainee Program. Jerome R. Bush. Prepared for the National Institute of Corrections in 1980.

This document was developed specifically for data collections arising out of a need to reduce jail crowding. It identifies the data elements that are of particular importance in understanding crowding. The section that describes charting the flow of the criminal justice system will be a useful tool for everyone who needs to clarify who has the authority to make which decisions in the system. This is a critical part of understanding how alternatives to incarceration "fit" into the criminal justice system. Data collection sheets and code books are included.

How to Handle Seasonality: Introduction to the Detection and Analysis of Seasonal Fluctuation in Criminal Justice Time Series Analysis. Carolyn Rebecca Block. Published by the Illinois Criminal Justice Information Authority in 1983.

Seasonality normally becomes an issue only when time series analysis or population forecasting is being done. However, it is something that every jail administrator experiences. Sometimes the jail is fuller than average. Typically, there are seasonal effects (the jail is fuller certain times of the year) and also weekly effects. This article offers approaches to dealing with this problem statistically. Understanding seasonality can help to determine how much over the average capacity to plan for.

The Intermediate Sanctions Handbook: Experiences and Tools for Policy Makers. Peggy McGarry and Madeline M. Carter, Center for Effective Public Policy. Published by the National Institute of Corrections and the State Justice Institute in 1993.

This document focuses on alternatives to incarceration. Two chapters are particularly applicable. Chapter 10 discusses the process used in building an information system to monitor sentencing, including variables that should be collected. Chapter 12 discusses an analytical approach to targeting offenders for specific sanctions and how data is used in this process. For those who think of alternatives as being "outside" the jail, these chapters help make clear just how tied together jails and alternatives are.

Jail Overcrowding Project Information Requirements. Duane Baker and Associates. Published by the American Justice Institute, Sacramento, California, in 1979.

This manual provides advice on how to plan a data collection that deals with incarceration practices. It includes sample forms, procedures, and data display techniques.

Policy Making in Criminal Justice: The Use of Hard Data at Each Stage of the Policy/Planning Process. Tim Brennan and David Wells. Published by the Northpointe Institute of Public Management, Traverse City, Michigan, in 1991.

This paper shows how data can be used at each stage of the planning and policymaking process. The examples used to illustrate each stage come from varied areas of the criminal justice system and from actual clients. This is a good "real-world" example of how data can be used.

Research Methods in Criminal Justice and Criminology. Frank E. Hagan. Published by Macmillan Publishing Company, New York, in 1982.

This document includes sections on the use of nationally available crime statistics, various survey methods (including victim surveys), the use of unobtrusive measures, and data collection strategies. It is a valuable resource.

Statistical Applications in Criminal Justice. Gennaro F. Vito and Edward J. Latessa. Published by Sage Publications, Newbury Park, California, 1989. This is Volume 10 of the Law and Criminal Justice Series.

This basic statistics text has been used in a variety of criminal justice educational settings. Its advantage over other basic statistical texts is that all of the examples used are from criminal justice.

Statistical Sampling Methods for Correctional Planners. Edward Lakner. Prepared by the National Clearinghouse for Criminal Justice Planning and Architecture. Published by the University of Illinois at Urbana-Champaign in 1976.

This book was written for correctional planners and uses very technical language, which makes it extremely difficult reading for anyone without a background in statistics. Despite this difficulty, it is the only resource found that deals with problems specific to using samples to estimate jail population statistics. For those who will be working with professional planners on data collections, it will be very helpful. Those who need more help to develop a sample than is available through this manual might be better off consulting one of the basic statistics books mentioned below.

Other Resources

Evaluator's Handbook. Joan L. Herman, Lynn Lyons Moms, and Carol Taylor Fitz-Gibbon. Published by Sage Publications, Beverly Hills, California, in 1987.

This handbook is one of a series of eight books called the ***Program Evaluation Kit***. It is very practical; takes a systematic, step-by-step approach to evaluation; and is written in everyday language. This volume explains what program evaluation is, differentiates between several kinds of evaluations, and identifies the steps in each. It provides many worksheets and similar tools that can be used in an evaluation. This is an excellent resource for those interested in evaluating jail programs and services.

Handbook in Research and Evaluation. Stephen Isaac and William B. Michael. Published by Edits Publishers, San Diego, California, in 1971.

This handbook is one of the most frequently used texts in research methods courses. It is divided into short sections about specific research issues, which makes it easy to find help in a specific area. The book is not as helpful as a guide on how to do research and, because it is written in rather technical language, a lay person may find it hard to understand. Other resources in this area may be more helpful for people not very familiar with research methods.

How to Analyze Data. Carol Taylor Fitz-Gibbon and Lynn Lyons Moms. Published by Sage Publications, Newbury Park, California, in 1987.

This book, part of the ***Program Evaluation Kit***, is used in NIC's program evaluation seminar as resource material. It is a basic introduction to a variety of elementary statistical techniques, including those for summarizing data, for examining differences between groups, and for examining relationships between two measures. Analysis of Effect Size—a relatively recent and simple approach for examining differences between groups and for conducting meta-analysis is shown. Only the most basic and useful statistical techniques that are appropriate for answering essential evaluation questions have been included. Worksheets and practical examples are given throughout the volume to support the use of each technique. Guidance is provided on the use of statistical techniques for constructing tests and questionnaires, on methods for choosing appropriate statistics, on using meta-analytic techniques, and on using statistical packages—particularly SPSSX (a statistical package for microcomputers)—giving readers all the information needed to properly analyze their data.

How to Calculate Statistics. Carol Taylor Fitz-Gibbon and Lynn Lyons Moms. Published by Sage Publications, Beverly Hills, California, in 1978.

This book is also part of the ***Program Evaluation Kit***. Plainly written, it takes the reader through calculating and displaying most of the statistics needed to analyze jail data. Worksheets are provided to use in calculations, which make the calculations more understandable. The accompanying text addresses most of the questions that statistical beginners might have.

How to Lie with Statistics. Darrell Huff. Published by W.W. Norton Company, New York, in 1954.

Although this book is more than four decades old, it still does an excellent job of identifying common statistical misuses, abuses, and misrepresentations. It identifies actions that statistical consumers can take to make sure they are not "duped" by statistics. This book is easy to read and informative.

Practical Research: Planning and Design. Paul D. Leedy. Published by Macmillan Publishing Company, New York, in 1974.

For those working with research methods for the first time, this manual would be a good investment. It is relatively jargon free and addresses research planning (which is an often underplayed but critical element), different research designs, and ways to present the results. This manual does a good job in providing guidelines for writing a research report. Illustrations and examples are used to show what research materials look like.

Social Statistics. (2nd edition.) Hubert M. Blalock, Jr. Published by the McGraw-Hill Book Company, New York, in 1979.

One of the most frequently used statistics texts, this book is sometimes difficult to read, primarily because of the highly technical language. If the reader is already familiar with statistics, it would be an excellent resource because of the exceptional detail included.

Statistics: The Essentials for Research. (3rd edition.) Henry E. Klugh. Published by Lawrence Erlbaum Associates, Hillsdale, New Jersey, in 1986.

This is another useful statistical text, with a good section on displaying data.

"Uncertainty, Policy Analysis and Statistics." James S. Hodges. Printed in **Statistical Science**, Volume 2, No. 3, August 1987, pp. 259–291.

This very technical paper deals with an important issue—uncertainty. Each of the statistical techniques has uncertainty attached to it. The academic world learns to live with uncertainty, but uncertainty presents problems in the politicized arena of criminal justice policy. This paper argues for constructing systems to eliminate that bias.

Understanding Statistics. (2nd edition.) William Mendenhall and Lyman Ott. Published by Duxbury Press, North Scituate, Massachusetts, in 1976.

For someone inexperienced in statistics, this book would be a good resource. It is readable, uses many examples, and is less confusing than many traditional statistics texts. Because it is basic, it does not have the detail that other texts might offer, which may be a bonus for beginners.

Appendix C: Manual Data Collection Procedures and Sample Forms

The manual information system presented in this appendix was used in a fairly large facility (100 beds, about 6,000 bookings per year) for 6 years. It provides basic information to management, is relatively easy to maintain, and does not demand much time on the part of clerical staff. And it works!

Elements of the System

To set up this management information system, a jail needs:

1. A supply of 5" x 8" index cards, preferably pre-printed with the names of the data elements to be collected and spaces for information to be recorded (@ about \$15 for 500, plus the cost of printing),
2. Two sets of alphabetic file dividers (@ about \$3 each),
3. Two card files (@ about \$2 each),
4. A ball point pen (@ \$.36), and
5. About 70 hours of a clerk's time for every 1,000 bookings. This translates to about 4 minutes for each card filled out.

So, if the jail booked 1,000 people in a year, the cost of maintaining this manual information system would be about \$40.36 for equipment (assuming you chose the least expensive equipment) and \$560 in direct labor costs (if the person maintaining the system makes \$8 an hour). At \$600.36 for the annual operating cost, this manual information system is within reach of any facility.

System Start-Up

For the discussion on identifying information that should be collected, refer to Chapter 2.

Step 1: Identify the information to be collected.

The purposes of this information system are to:

- Collect the Inmate Population Data Elements (Average Daily Population; Length of Stay; Jail Days; Total Bookings; Net Bookings; Number of Felons, Misdemeanants, Traffic Offenders, and

Holds; Number of Pre-Trial, Post-Trial, and Sentenced Inmates; Number of Male and Female Inmates; and Adult and Juvenile Inmates).

Collect some basic Inmate Profile Data Elements (Charge, Age, Ethnicity, Residence, and Release Status).

- Collect some Operational Data Elements (Number and Location of Transports; Number of Incidents; and Number of Contact, Non-Contact, Professional, and Family Visits).

Step 2: Determine how frequently summary reports are needed,

Most sheriffs and jail administrators need monthly statistical reports. The Inmate Population Data Elements are most understandable statistically when **grouped** by month—weekly statistics are too much to absorb when looking for jail population trends and patterns. An exception is when the jail is trying to control crowding—monthly statistical reports would not be frequent enough.

Step 3: Design an inmate information card.

Figure C.1 provides a sample Inmate Information Card

Step 4: Decide who will maintain the system.

One person, preferably a clerk, should be responsible for maintaining the information system. If the jail has no clerical staff, officers on the graveyard shift or dispatchers might be able to maintain the system during slow periods. One benefit to having clerical staff maintain the information system is that they are not tied to radios as dispatchers are and can leave their work area to get the information.

Step 5: Determine where the jail captures the information and how to get it from that source to the card.

It is hard to be specific about this step because jails capture the information in so many places. The Daily Activity Log or a Shift Activity Sheet, Booking Log, and the Arrest Report probably provide most of the Inmate Population and Inmate Profile Data Elements. The Operational Data Elements may come from the Transport, Incident, and Visitor Logs. Another potential resource is the status board found in the Booking Room or in Central Control, which records who is in custody, some descriptive information, and their housing assignments.

Assume, for this example, that the inmate information can be collected from a Daily Activity Log or Shift Activity Sheet (Figure C.2). (Legal size paper could be used if activity warrants.)

Assume also that none of the logs needed to gather the operating information about transports, incidents, and visitors is available. This will require some additional items:

- Three loose-leaf binders, and
- A ream of copy paper.

Repeat Steps 1 through 4 to decide:

- What information should be included in these logs;
- How frequently the information will be needed;
- The design of the Visitor, Transport, and Incident Log Sheets; and
- Who will maintain them.

Figures C.3, C.4, and C.5 provide sample Visitor, Transport, and Incident Logs.

Figure C.1 Sample Inmate Information Card

Inmate Information Card			
Name: _____	Book-In Date: _____	Book-Out Date: _____	
Birth Date: _____	Race: _____	Gender: _____	Length of Stay: _____
Charges: _____			
<input type="checkbox"/> Felony	<input type="checkbox"/> Pre-Trial	<input type="checkbox"/> Released at Booking	
<input type="checkbox"/> Misdemeanor	<input type="checkbox"/> Post-Trial	<input type="checkbox"/> Housed	
<input type="checkbox"/> Traffic	<input type="checkbox"/> Sentenced	<input type="checkbox"/> Hold for:	_____
<input type="checkbox"/> Other:	_____		
<input type="checkbox"/> County Resident	Released To: _____		
<input type="checkbox"/> State Resident	Bond Type: _____		
<input type="checkbox"/> Out-of-State Resident	Bond Amount: _____		
<input type="checkbox"/> Undocumented Alien			
Employed: <input type="checkbox"/> Full-Time	<input type="checkbox"/> Part-Time	<input type="checkbox"/> Unemployed	<input type="checkbox"/> Student

Figure C.2 Daily Activity Log

Any County Sheriffs Department Jail Division Daily Activity Log		Date _____	Shift _____						
IN:		OUT:							
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Time:	_____	Time:	_____						
Name:	_____	Name:	_____						
Charge:	_____	Release to:	_____						
P.D.:	_____	Bond:	_____						
Statistical Summary: Beginning Count _____ End Count _____									
# Felons	_____	# Misdemeanants	_____	# Traffic Offenders	_____	# Other	_____		
# Pre-Trial	_____	# Sentenced	_____	# Females	_____	# Juveniles (Male)	_____	(Female)	_____

Figure C.3 Sample Visitor Log

Figure C.4 Sample Transport Log

Any County Sheriffs Department
Jail Division
Transport Log

Officer Name:

- | | | |
|----------|---------------------|-----------------|
| 1. _____ | Regular Hours _____ | Comp/O.T. _____ |
| 2. _____ | Regular Hours _____ | Comp/O.T. _____ |
| 3. _____ | Regular Hours _____ | Comp/O.T. _____ |
| 4. _____ | Regular Hours _____ | Comp/O.T. _____ |
| 5. _____ | Regular Hours _____ | Comp/O.T. _____ |

Date of Transport: _____

Reason for Transport: _____

Inmate Name:

- | | | |
|----------|-------------|-----------|
| 1. _____ | From: _____ | To: _____ |
| 2. _____ | _____ | _____ |
| 3. _____ | _____ | _____ |
| 4. _____ | _____ | _____ |
| 5. _____ | _____ | _____ |
| 6. _____ | _____ | _____ |
| 7. _____ | _____ | _____ |
| 8. _____ | _____ | _____ |

Time of Departure: _____

Meal Cost:

Time of Return: _____

1. \$ _____

2. \$ _____

Total \$ _____

Vehicle:

Starting Mileage: _____ Vehicle Used: _____

Ending Mileage: _____ Condition: _____

Total Mileage: _____

Gas Cost: _____ Gallons Purchased: _____

Related Transport Problems: _____

Submitted: _____ Date: _____

Figure C.5 Sample Incident Log

Any County Sheriffs Department
Jail Division
Facility Incident Log

Background:

Inmate Name: _____ Booking Number: _____

Housing Assignment: _____ Cell Assignment: _____

Incident Date: _____ Time: _____ Reporting Employee: _____

Report Date: _____ Time: _____ Supervisor: _____

Location:

<input type="checkbox"/> Holding Cell	<input type="checkbox"/> Comdor	<input type="checkbox"/> Security Visiting
<input type="checkbox"/> Booking Room	<input type="checkbox"/> Multi-Use Room	<input type="checkbox"/> Contact Visiting
<input type="checkbox"/> Vehicle Sallyport	<input type="checkbox"/> Indoor Exercise	<input type="checkbox"/> Housing Unit #: _____
<input type="checkbox"/> Isolation Cell	<input type="checkbox"/> Outdoor Exercise	<input type="checkbox"/> Cell #: _____
<input type="checkbox"/> Other: _____		

Reason:

<input type="checkbox"/> Suicidal	<input type="checkbox"/> Causing a Disturbance	<input type="checkbox"/> Contraband
<input type="checkbox"/> Assault on Staff	<input type="checkbox"/> Threatening Staff	<input type="checkbox"/> Escape Attempt
<input type="checkbox"/> Assault on Inmate	<input type="checkbox"/> Threatening Inmate	<input type="checkbox"/> Escape
<input type="checkbox"/> Illness	<input type="checkbox"/> Arson	<input type="checkbox"/> Destroy Jail Property
<input type="checkbox"/> Injury	<input type="checkbox"/> Protective Custody	<input type="checkbox"/> Refuse an Order
<input type="checkbox"/> Other: _____		

Description of Incident:

Rule Violation: Yes Major Minor No New Offense: Yes No

Disciplinary Action:

Major Violation:	Minor Violation:
<input type="checkbox"/> Immediate segregation	<input type="checkbox"/> No action
<input type="checkbox"/> Move to more restrictive housing	<input type="checkbox"/> Verbal reprimand
<input type="checkbox"/> Short-term lockdown	<input type="checkbox"/> 48-hour loss of privilege: _____
	<input type="checkbox"/> 24-hour lockdown
	<input type="checkbox"/> Other: _____

Use of Restraints/Use of Force:

<input type="checkbox"/> None	<input type="checkbox"/> Staff	<input type="checkbox"/> Inmate
<input type="checkbox"/> Physical force/restraint	<input type="checkbox"/> Facility Property	<input type="checkbox"/> Staff Property
<input type="checkbox"/> Mechanical restraint	<input type="checkbox"/> Inmate Property	
<input type="checkbox"/> Intermediate weapon		

Disciplinary Hearing Required: Yes No

Hearing Date: _____ Date Served: _____

Grievance Filed: Yes No Additional Comment: _____

Administrative Review and Approval:

Detention Officer: _____	Date: _____
Supervisor: _____	Date: _____
Jail Administrator: _____	Date: _____

The easiest ways to get this information are to:

- Require that all incidents in the facility be documented on the Incident Log. To do that, the sheriff or jail administrator must define for staff what kind of events must be documented. They should review each form and then give it to the person who is maintaining the information system to be placed in the proper binder.
 - Require that all officers transporting inmates outside the facility begin a Transport Log prior to departing the jail and complete it as **soon** as they return. Have them give the completed form to the person who is maintaining the system to be placed in the proper binder.
- Require that all persons who visit the facility complete a Visitor Log entry. The forms should be pre-punched and placed in a three-ring binder, kept at the visitors' entrance.

If the jail has more than one source for a particular data element, a decision must be made about which source to use. Two conditions might help with that choice:

1. One source may be more accurate than another.
2. The fewer sources providing the information, the better. If one source has only one data element and another source has more than one, unless there is a **compelling** reason to use the former, use the source that provides the most data elements.

Step 6: Train the person who will maintain the system in how to fill out the inmate information card. Insist that it be done daily.

Information systems are manageable if they are maintained daily, but are extremely tedious to update. Worse, any missing information may be impossible to get once the inmate is gone and the staff's memories have faded. Occasionally, it is also a good idea to make sure the information filled out on each of the Inmate Information Cards is accurate. Be sure to explain how each data element is to be entered and review the **rules** for coding data in **Section 4**.

Step 7: Pick a day to start.

Ideally, the information system should start on the first of the month. Allow a few days of practice to be sure that the system is working properly. On Day 1, make a card for each inmate who was in custody at 8:00 a.m. Label two card files: "Current Inmates" and "Past Inmates." File the Inmate Information Cards by inmate last name after the appropriate letter in the Current Inmate Card File. Get a calendar small enough to fit inside one of the card files. At the same time each day, record the number of inmates in custody. Be sure to pick a time when the jail's population is not artificially low (like right after court). Usually 8:00 a.m. is a good time.

Step 8: Begin daily maintenance.

Each day, complete an Inmate Information Card for every person booked at the jail. Pull out the Inmate Information Card of every person who is released, enter the release information, and file it in the Past Inmate Card file.

Step 9: Prepare the first monthly report.

Begin preparing a monthly report by averaging the daily head counts to calculate the jail's average daily population. Before starting the monthly report, make "tally sheets" for each data element. Figure C.6 provides a sample of a monthly tally sheet for the data element "Inmate Age."

Take out the Past Inmate Card file and remove all the Inmate Information Cards. Be careful to use only the Past Inmate Card file. The system actually counts inmates as they are released from the jail rather than counting them as they are booked. Statistically this is better, because it captures information sooner about people who have long lengths of stay. Since inmate turnover in jails is relatively rapid, it will take only about a month for the system to include people who have been in jail for periods longer than a month. If information is gathered on admission, it will take many months, perhaps more than a year, to include people who stay very long periods of time. This avoids sample bias.

When going through the Inmate Information Cards, make a mark on the section of a tally sheet that represents the value noted on the card. For instance, put a check mark next to "Felony" if the prisoner is charged with a felony. If there is more than one charge, always code the most serious charge. When finished with the inmate information, pull the Transport, Incident, and Visitor Logs for the month and follow the same process. Summarize the information in report form. For some help in doing this, refer to Chapter 8 for some tips on how to display data.

Next, make an accounting sheet for each of the data elements so when each monthly report is completed, a "Year-to-Date" total can be calculated. This will make doing an annual report easier. Put these cards together and label them by month. When the annual report is completed, store the cards in a dry, secure place. They may be useful in later data collections.

Step 10. Evaluate the information system in 6 months and make any necessary changes.

Every jail has its own peculiarities and this system may not exactly fit your facility. After six months, however, you should have enough experience with it to know what portions should be changed or adapted to better fit your needs. Set a time to do this. Be sure to talk with the person who is maintaining the system. This information system, when tailored to meet the needs of the jail, can be a powerful management tool. Take the time to improve on it!

Figure C.6 Sample Monthly Tally Sheet for Inmate Age

**Any County Sheriffs Department
Jail Division
Inmate Information System Monthly Tally Sheet**

Age	Male	Female	Age	Male	Female
18			34		
19			35		
20			36		
21			37		
22			38		
23			39		
24			40		
25			41		
26			42		
27			43		
28			44		
29			45		
30			46		
31			47		
32			48		
33			49 or older		
# of Males					
# of Females					

Appendix D: Inmate Profile Data Collection

Data Collection Form

Inmate Name: _____

ID #	_____	Arresting Agency	_____
Booking #	_____	Occupation	_____
Residence	_____	Employment Status	_____
Date of Birth	_____/_____/_____	Marital Status	_____
Race	_____	Booking Date	_____/____/_
Legal Status	_____	Booking Time	____:____
Hold Type	_____	Release Date	_____/____/_
Charge Type	_____	Release Time	____:____
Charge 1	____	Release Status	____
Charge 2	____	Bond Type	____
Charge 3	____	Last Grade Attended	____
Charge 4	____	Military Status	_____
Total # Charges	____	Security Status	_____

Priors (Felony)	■ ■	Does the subject have health problems that require immediate treatment?	■
Priors (Misdemeanor)	■ ■	Is the subject under the influence of drugs or alcohol?	■
Convictions	■ ■	Does the subject exhibit psychological problems that require immediate treatment?	■
Prior Violent	■ ■	Is the subject currently under a doctor's care?	■
Prior Drug	■ ■	Is the subject currently receiving psychiatric care?	■
Prior Alcohol	■ ■	Has the subject ever been in treatment for mental health issues?	■
Days Served	■ ■ ■ ■	Treatment type.	■
Disposition Date	■ ■ / ■ ■ / ■ ■	Does the subject have a history of substance abuse?	■
Sentencing Date	■ ■ / ■ ■ / ■ ■	Primary substance.	■
Continuances	■ ■	Has the subject ever been in treatment for substance abuse?	■
Court	■ ■	Treatment type.	■
Judge	■ ■		

Inmate Profile Data Collection Code Book

Variable Name	Codes	Explanation
ID Number		A three-digit number assigned to each case on which data is collected (e.g., 001,002,003, etc.).
Booking Number		A number assigned to each case booking. Note: The number of digits will vary based on the booking number protocol.
Residence	1 2 3 4 5 6 8 9	City of _____ City of _____ City of _____ Other in-county In-state, out of county Out of state Out of U.S. Unknown
Sex	1 2	Male Female
Date of Birth		A six-digit number (e.g., 01/04/76)
Race	1 2 3 4 5 6	Caucasian African-American Hispanic Native American Asian Other
Legal Status	1 2 3 4	Pre-trial Sentenced Hold only Other
Hold Type	1 2 3 4 5	Hold for other county or state Hold for parole/probation Hold for immigration Hold for AWOL Hold for _____
Charge Type	1 2 3 4	Felony Misdemeanor Traffic Other

Variable Name	Codes	Explanation
Charge 1-4 (Person Offense Codes)	10 11 12 13 14 15 16 17 18 19	Murder Attempted murder Manslaughter Aggravated assault Robbery, armed Robbery, unarmed Minor assault Kidnapping Menacing Other offenses against persons
Charge 1-4 (Property Offense Codes)	20 21 22 23 24 25 26 27 28 29	Larceny Auto theft Burglary Breaking and entering Arson Theft Shoplifting Criminal mischief Trespassing Other offenses against property
Charge 1-4 (Offenses Against Family or Children Codes)	30 31 32 33 34 35 36 37 38 39	Non-support Failure to provide Desertion Neglect Bigamy Adultery Contributing to the delinquency of a minor Violation of compulsory school law Paternity offenses Other offenses against family or children
Charge 1-4 (Sex Offenses Codes)	40 41 42 43 44 45 46 47 48 49	Forcible sex acts Unnatural sex acts Prostitution Immoral acts Promiscuity Illegal sexual relations Pornography Soliciting Pandering Other sex offenses

Variable Name	Codes	Explanation
Charge 1-4 (Crimes of Forgery, Fraud, and Conspiracy Codes)	50 51 52 53 54 55 56 57 58 59	Forgery Fraud Deception/embezzlement Uttering fraudulent instrument Issuing fraudulent instrument Conspiracy Blackmail, extortion Receiving/concealing stolen property Impersonation Other crimes of forgery , fraud, and conspiracy
Charge 1-4 (Drug, Alcohol, and Weapon Offenses Codes)	60 61 62 63 64 65 66 69	Violation of weapons laws Violation of liquor laws Violation of narcotics laws Violation of gambling laws Drunk or drinking City ordinance violations (except disorderly conduct) Disorderly conduct Other weapon, drug, or alcohol offenses
Charge 1-4 (Traffic Violations Codes)	70 71 72 73 74 79	Moving violation Standing violation Operating a motor vehicle without license or registration Operating a motor vehicle under the influence of drugs or alcohol Careless or reckless driving Other traffic or motor vehicle offenses
Charge 1-4 (Miscellaneous Offenses Codes)	80 81 82 83 84 85 86 87 88 89 90 91 92 93 99	Escape from custody Obstructing Habitual criminal Harboring a fugitive Cruelty to animals Violation of court order FTA/FTP/bond revocations Appeals Writs Temporary holds Probation violation Parole violation Protective custody Other: _____ Unknown

Variable Name	Codes	Explanation
Total Number of Charges		Total number of charges this booking.
Arresting Agency	1 2 3 4 5 6 7 8 9	County of _____ City of _____ City of _____ City of _____ State police or patrol Other state agency Federal Other county Unknown
Occupation	1 2 3 4 5 6 7 8 9 10 11 12 99	Professional, technical Managerial, administrative Sales Clerical Craftsman Operator (e.g., machine operator or factory worker) Transportation equipment operator Laborer (except farm) Farm laborer Service worker Student No employment history Unknown
Employment Status	1 2 3 4	Full-time employed Part-time employed Unemployed Never employed
Marital Status	1 2 3 4 5 6 9	Married Single Divorced Separated Common-law Widow(er) Unknown
Booking Date		Six-digit date of booking (e.g., 04/19/96).
Booking Time		Time of booking in military time (e.g., 04:25, 17:45).

Variable Name	Codes	Explanation
Release Date		Six-digit date of release (e.g., 09130197).
Release Time		Time of release in military time (e.g., 23:45).
Release Status	01 02 03 04 05 06 07 08 09 10 11 12 13 14	Jail sentence completed Acquitted, charges dropped Transferred to state or federal institution Placed on probation, probation reinstated Paroled, parole reinstated Bail or bond Release to other jurisdiction Release to mental health facility Release to medical facility Release to treatment program/facility Deferred sentence or prosecution Escape or walk-away Fine paid Other
Bond Type	01 02 03 04 05 06 07 08 09 10 11 12 13 14	Personal recognizance —pre-trial services Personal recognizance —court ordered Cash Surety Property 10% bond Personal recognizance with cosigner(s) Personal recognizance with conditions Personal recognizance and surety Personal recognizance and cash Personal recognizance and property Cash and surety Bond reinstated Other:
Last Grade		A two-digit number that indicates the last grade completed (e.g., 13 = one year of college).
Military Status	1 2	Veteran Non-veteran

Variable Name	Codes	Explanation
Security Classification	1 2 3 4 8	Maximum Medium Minimum Special housing Not applicable—another data numbering convention
Prior Arrests (Felony)		A two-digit number that indicates the number of prior felony arrests in this county.
Prior Arrests (Misdemeanor)		A two-digit number that indicates the number of prior misdemeanor arrests in this county.
Number of Convictions		A two-digit number that indicates the number of convictions in this county.
Number of Violent Arrests		A two-digit number that indicates the number of arrests in this county for offenses defined as violent. Note: Each jurisdiction needs to identify the charges it includes in this category.
Number of Drug Arrests		A two-digit number that indicates the number of arrests in this county for offenses involving drugs.
Number of Alcohol Arrests		A two-digit number that indicates the number of arrests in this county for offenses involving alcohol.
Days Served		A computed variable, the number of days between date of booking and date of release. Note: If less than 24 hours, express as a decimal (portion of a day), e.g., 12 hours = .5
Date of Disposition		A six-digit date that shows the date trial was held, plea was entered, or case was dismissed.
Date of Sentencing		A six-digit date that shows the date of sentencing.
Number of Continuances		A two-digit number that shows the number of continuances granted prior to trial on this charge.
Court	1 2 3 4 5	Superior court Lower court Municipal court Appeals court Other _____
Judge	1 2 3	_____

Variable Name	Codes	Explanation
Health Problems	1 2	Yes No
Influence of Drugs, Alcohol	1 2	Yes No
Mental Health Problems	1 2	Yes No
Doctor's Care	1 2	Yes No
Current Mental Health Care	1 2	Yes No
Previous Mental Health Care	1 2	Yes No
Treatment Type	1 2 3	Inpatient Outpatient Other
Substance Abuse History	1 2	Yes No
Primary Substance	1 2 3 4 5 6 7 8 9	Alcohol Marijuana Hallucinogens Amphetamines Cocaine Heroin, other opiates Crack or crank PCP Other
Substance Abuse Treat- ment	1 2	Yes No
Treatment Type	1 2 3	Inpatient Outpatient Other

Appendix E: Incident Data Code Book

Variable	Code	Explanation
Housing Assignment	11	Housing Unit # _____
	12	Housing Unit # _____
	13	Housing Unit # _____
	14	Housing Unit # _____
	15	Housing Unit # _____
	99	Not yet assigned to housing
Location of Incident	01	Holding
	02	Booking
	03	Intake sallyport
	04	Property room, search/shower
	05	Isolation
	06	Corridor
	07	Library or classroom
	08	Exercise
	09	Visiting
	10	Kitchen
	11	Housing Unit # same as above _____
	12	Housing Unit # same as above _____
	13	Housing Unit # same as above _____
	14	Housing Unit # Same as above _____
	15	Housing Unit # same as above _____
	16	Other (in facility)
	17	Outside facility
	88	Not applicable
	99	Unknown

Variable	Code	Explanation
Reason	01 02 03 04 05 06 07 08 09 10 11 12 13 14 88 99	Suicidal Violent Illness Accidental injury Causing a disturbance Harming/threatening staff Harming/threatening inmate Arson Protective custody Contraband violation Escape attempt/risk Refusal to obey an order Destruction of jail property Other Not applicable Unknown
Rule Violation	1 2 3	Major Minor None
Crime Report Required	1 2	Yes No
Major Violation	1 2 3 4 8 9	Move to isolation Move to maximum security Move to special management Move to other Not applicable Unknown
Minor Violation	1 2 3 4 5 6 8 9	No action Verbal reprimand 48-hour loss of privilege 24-hour lockdown Over 48-hour loss of privilege Other _____ Not applicable Unknown

Variable	Code	Explanation
Privilege Loss	1 2 3 4 5 6 7 8 9	Contact visit Exercise Library Program participation _____ Television Non-attorney telephone Other Not applicable Unknown
Use of Restraints	1 2 3 8 9	None needed Physical Mechanical Not applicable Unknown
Injury To	1 2 3 8 9	Staff Inmate Other Not applicable Unknown
Review Board	1 2 8 9	Yes No Not applicable Unknown
Date of Review		A six-digit date that documents the date of the review board hearing.
Date Notice Served		A six-digit date that documents the date the inmate was served notice of the review board hearing.
Grievance Filed	1 2 8 9	Yes No Not applicable Unknown

Appendix F: Transport Data Collection

Variable	Code	Explanation
Reason for Transport	01	Emergency room treatment
	02	Scheduled medical, dental, or other treatment appointment
	03	Court appearance
	04	Pick-up in another county on warrant
	05	Return from this county to another county on its warrant
	06	Transport to state Department of Corrections
	07	Transport to state hospital
	08	Transport to treatment facility
	09	Court ordered
	10	Other
	88	Not applicable
	99	Unknown

Variable	Code	Explanation
Location	01	Hospital
	02	Doctor's office
	03	Dental office
	04	Psychiatrist/psychologist office
	05	Mental health clinic
	06	Public health department
	07	Other health-related transport
	08	State prison
	09	State reformatory
	10	State hospital
	11	Veteran's hospital
	12	Halfway house
	13	Other treatment placement
	14	County _____
	15	County _____
	16	Other county
	50	Civil court
	51	District court
	52	County court
	53	Juvenile court
	54	Municipal court
	55	Other court
	88	Not applicable
	98	County jail
	99	Unknown
Departure Time		A four-digit number on the 24-hour (military) clock that designates the time officers left the jail for the transport .
Return Time		A four-digit number on the 24-hour (military) clock that designates the time officers returned to the jail.
OT/Comp	1	Yes
	2	No
	8	Not applicable
	9	Unknown
Meals		A four-digit number that indicates in dollars and cents the amount of money spent for food during the transport.
Odometer Start		A six-digit number that records the odometer reading on the vehicle at departure.

Variable	Code	Explanation
Odometer End		A six-digit number that records the odometer reading on the vehicle on return from the transport.
Type of Vehicle	1 2 3 8 9	Van Car Other Not applicable Unknown
Vehicle Number		A two-digit number that documents the fleet vehicle number of the vehicle used on the transport.
Cleanliness of Vehicle	1 2 3	Poor Fair Good
Mechanical Condition of Vehicle	1 2 3	Poor Fair Good
Gas Cost		A four-digit number that documents in dollars and cents the amount of money spent to purchase gasoline on the transport.
Number of Gallons		A two-digit number that documents the number of gallons of gasoline that were purchased on the transport.

Appendix G: Tables for Determining Sample Size

This section is reprinted from *Statistical Sample Methods for Correctional Planners*, by Edward Lakner, prepared at the National Clearinghouse for Criminal Justice Planning and Architecture and published by the University of Illinois at Urbana-Champaign in 1976, pp. 68–71.

Tables G.1 through G.5 give the sample sizes needed to estimate proportions in selected populations at the 95% and 99% confidence levels, with allowable errors of $\pm 2\%$, 3%, and 5%. By interpolating between the marginal¹ values of the table (the expected rate of occurrence and the size of the population), approximate intermediate sample sizes may be found. Exact sample sizes for proportions and population sizes not listed in the tables can be found by using the formulas provided in Chapter 5.

Table G.1 Sample Size for Reliability of $\pm 2\%$ at the 95% Confidence Interval

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	239	316	377	400	410	413
1,000	314	464	606	669	697	706
1,500	350	549	758	859	907	922
2,000	372	604	869	1,004	1,071	1,091
2,500	386	642	952	1,117	1,200	1,225
10,000	436	796	1,332	1,678	1,873	1,936
50,000	452	850	1,491	1,939	2,203	2,291
100,000	456	863	1,532	2,009	2,294	2,390

¹Marginals in this sense mean the numbers around the outside of the table.

Table G.2 Sample Size for Reliability of $\pm 3\%$ at the 95% Confidence Interval

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	144	217	288	320	336	340
1,000	169	278	406	473	506	551
1,500	179	306	470	561	609	624
2,000	184	322	509	619	678	696
2,500	188	333	537	660	727	748
10,000	199	370	639	823	929	964
50,000	202	381	674	881	1,004	1,045
100,000	203	384	682	895	1,023	1,065

Table G.3 Sample Size for Reliability of $\pm 5\%$ at the 95% Confidence Interval

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	64	108	165	196	213	217
1,000	68	121	198	244	262	278
1,500	70	126	211	269	297	306
2,000	70	130	219	278	312	322
2,500	71	131	224	286	322	333
10,000	72	136	240	313	356	370
50,000	72	137	245	321	367	380
100,000	73	138	246	322	369	384

Table G.4 Sample Size for Reliability of $\pm 2\%$ at the 99% Confidence Interval

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	305	374	420	437	444	446
1,000	441	598	725	776	798	804
1,500	517	748	957	1,047	1,087	1,099
2,000	565	855	1,138	1,267	1,328	1,346
2,500	599	935	1,284	1,451	1,531	1,901
10,000	731	1,299	2,098	2,584	2,848	2,932
50,000	776	1,450	2,520	3,257	3,688	3,830
100,000	787	1,489	2,640	3,460	3,950	4,113

Table G.5 Sample Size for Reliability of $\pm 3\%$ at the 99% Confidence Interval

Population Size	5% or 95%	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Expected Rate of Occurrence in the Population						
500	206	285	351	378	390	393
1,000	260	399	541	607	638	647
1,500	284	460	661	761	788	826
2,000	298	498	742	873	939	959
2,500	307	525	801	955	1,036	1,061
10,000	338	622	1,055	1,341	1,504	1,556
50,000	348	655	1,152	1,502	1,709	1,778
100,000	350	663	1,177	1,544	1,763	1,836

Appendix H: Simple Random Sampling

This appendix is taken directly from *Practical Research; Planning and Design*, by Paul D. Leedy, published by Macmillan Publishers, 1974, pp. 93–99. Simple random sampling theory and practice is described in considerable detail.

Types, Methods, and Procedures of Sampling

One basic rule holds whenever a researcher is considering methodology in relation to data—it matters not whether this methodology concerns sampling, a statistical procedure, or any other type of operation, and that one basic rule is: *Look carefully at the nature, the characteristics, the quality of the data.* After the researcher sees this clearly, he/she can then more intelligently select the proper methodology for the treatment of those data.

The composition of the sample is derived by selecting the sample units from those of a much larger population. In survey studies the *manner* in which the sample units are selected is very important. Generally, the components of the sample are chosen from the population universe by a process known as *randomization*. Such a sample is a *random sample*. The two elements that are more important than any others in survey research are randomization and bias. We shall discuss randomization here....

Randomization means *selecting apart of the wholepopulation in such a way that the characteristics of each of the units of the sample approximate the broad characteristics inherent in the total population.* Let's explain that. I have a beaker which contains 100 cc. of water. I have in another container a concentrated solution of 10 cc. of acid. I combine the water and the acid in proportions of 10 to 1. After thoroughly mixing the water and the acid, I should be able to extract 1 cc. from any part of the solution and find in that 1 cc. sample a mixture of water and acid in precisely a 10:1 proportion. Just so, if we have a conglomerate population with variables such as differences of race, wealth, education, and other factors, and *if I have a perfectly selected random sample* (a situation that is more theoretical than practical), I will find in the sampling those same characteristics of race, wealth, education, and so forth that exist in the larger population universe.

A sample is no more *representative* of the total population, therefore, than the *validity* of the method of randomization employed in selecting it. There are, of course, many methods of random selection. We shall look at merely a few of the more common ones.

The Roulette Wheel Method

If the population is small—50 to 75 individuals—each individual may be assigned a number in some orderly sequence: alphabetically by surname, by birthdate (youngest to oldest or the reverse), by the

respective weight of each individual, or by any other systematic arrangement. Corresponding numbers are on a roulette wheel. A spin of the wheel and its fortuitous stopping at a particular number selects the individual assigned to that number as a unit of the sample. The process of spinning the wheel and selecting the sample unit goes on until all the individuals needed to compose the sample have been chosen.

The Lottery Method

In the lottery method the population again is arranged sequentially and assigned a numerical identification. Corresponding numbers are marked on separate tabs and put into a revolving drum or closed container. The numbers are tossed so that they are thoroughly intermixed. Then one tab bearing a number is selected from the total number of tabs in the container without the selector seeing the pool. The number is recorded and *the tab is then tossed back into the tab pool*. This is an important feature of the lottery method. It ensures that every individual has the same chance of being chosen as every other individual. If, for example, we are selecting 50 units out of a population of 100 and we do not cast each tab back after it has been selected, we will have an ever-diminishing population from which to make the choices. And whereas the first choice would have 1 in 100 chances of being selected, the last unit chosen would have 1 in 50 chances of being selected or, in other words, his chances of being selected would be twice as great as those of the first individual.

In the event of the same number being drawn twice, the second drawing is ignored; the number is returned to the pool; the entire mass of numbers is tumbled again, and another drawing is made. Drawing and tumbling go on until 50 tabs have been selected purely by chance.

The Table of Random Numbers Method

The table of random numbers method is perhaps the most frequently used method for the random selection of a sample. Following is a part of a table of random numbers. We may employ this table in **any** manner in which we choose to use it. Generally the researcher enters the table according to some predetermined method.

Entrance into the table may, in fact, be accomplished in many ways. One fundamental principle must be kept in mind, however: *The purpose of randomness is to permit blind chance to determine the outcomes of the selection process to as great a degree as possible*. Hence, in determining a starting point for the selection of random numbers, chance must be given free reign.

Consider the following tables of random numbers (Tables H.1-H.4). Considering the second table (Table **H.2**) as a continuation of the first table (Table H.1), and counting from left to right, there are five blocks of random numbers in the horizontal arrangement. Vertically there are 10 blocks **from** top to bottom of the table.

Table H.1 Random Number Table (#1)

9774246762	4281 145720	4253323732	2707360751	2451798973
16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59	88 97 54 14 10
12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21	88 26 49 81 76
03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 28 16 80 45	60 11 14 10 95
55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	08 18 44 32 53	23 83 01 30 30
16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	87 35 20 96 43	84 26 34 91 64
84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	21 76 33 50 25	83 92 12 06 16
63 01 63 78 59	18 95 55 67 19	98 10 50 71 75	12 86 73 58 07	44 39 52 38 79
33 21 12 34 29	78 64 58 07 82	52 42 07 44 38	15 51 00 13 42	99 66 02 79 54
57 60 86 32 44	09 47 27 96 54	49 17 48 09 62	90 52 84 77 27	08 02 73 43 28
18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	06 76 50 03 10	55 23 64 05 05
28 62 38 97 75	84 16 07 44 99	83 11 46 32 24	20 14 85 88 45	10 93 72 88 71
23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	32 98 94 07 72	93 85 79 10 75
62 38 28 19 95	50 92 26 11 97	00 56 76 31 38	80 22 02 53 53	88 60 42 04 53
37 85 94 35 12	83 39 50 08 30	42 34 07 96 88	54 42 06 87 98	35 85 29 48 39
70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	17 76 37 13 04	17 74 21 19 30
56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	70 33 24 03 54	97 77 46 44 80
99 49 57 22 77	88 42 95 45 72	16 64 36 18 00	04 43 18 66 79	94 77 24 21 90
18 08 15 04 72	33 27 14 34 09	45 59 34 68 49	12 72 07 34 45	99 27 72 95 14
31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	52 85 66 60 44	38 68 88 11 80
68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52	68 07 97 06 57
74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76	15 54 55 95 52
27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	96 46 92 42 45	97 60 49 04 91
00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 66 04 26	11 04 96 67 24
29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	34 25 20 57 27	40 48 73 51 92
18 90 82 66 59	83 62 64 11 12	67 19 00 71 74	60 47 21 29 68	02 02 37 03 31
11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	76 70 90 30 86	38 45 94 30 38
35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	16 92 53 56 16	02 75 50 95 98
38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	40 01 74 91 62	48 51 84 08 32
31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	00 52 43 48 85	27 55 26 89 62
66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	76 83 20 37 90	57 16 00 11 66
14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
20 46 78 73 90	97 51 40 14 02	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
64 19 58 97 79	15 06 15 93 20	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
05 26 93 70 60	22 35 86 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
70 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 05 77 09 51
68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43
17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	44 95 27 36 99	02 96 74 30 83
90 26 59 21 19	23 52 23 33 12	96 93 02 18 39	07 02 18 36 07	25 99 32 70 23
41 23 52 55 99	31 04 49 69 98	10 47 48 45 88	13 41 43 89 20	97 17 14 49 17
60 20 50 81 69	31 99 73 68 68	35 81 33 03 76	24 30 12 48 60	18 89 10 72 34
91 25 38 05 90	94 58 28 41 36	45 37 59 03 09	90 35 57 29 12	82 62 54 65 60
34 50 57 74 37	98 80 33 00 91	09 77 93 19 82	74 94 80 04 04	45 07 31 66 49
85 22 04 39 43	73 81 53 94 79	33 62 46 86 28	08 31 54 46 31	53 94 13 38 47
09 79 13 77 48	73 82 97 22 21	05 03 27 24 83	72 89 44 05 60	35 80 39 84 88
88 75 80 18 14	22 95 75 42 49	39 32 82 22 49	02 48 07 70 37	16 04 81 67 87
90 98 23 70 00	39 90 03 08 90	55 85 78 38 36	94 37 30 69 32	90 89 00 76 33

Table H.2 Random Number Table (#2)

63 38 06 86 54	99 00 65 26 94	09 82 90 23 07	79 62 67 80 60	75 91 12 81 19
35 30 58 21 46	06 72 17 10 94	25 21 31 75 96	49 28 24 00 49	55 65 79 78 07
63 43 36 8269	65 51 18 37 88	61 38 44 12 45	32 92 85 88 65	54 34 81 85 35
98 25 37 55 26	01 91 82 81 46	74 71 12 94 97	24 02 71 37 07	03 82 18 66 75
53 74 23 89 67	61 32 28 69 84	94 62 67 86 24	98 33 41 19 95	47 53 53 38 09
02 83 21 17 69	71 50 80 89 56	38 15 70 11 48	43 40 45 86 98	00 83 26 91 03
64 55 22 21 82	48 22 28 06 00	61 54 13 43 91	82 78 12 23 29	06 66 24 12 27
85 07 26 13 89	01 10 07 82 04	59 63 69 36 03	69 11 15 83 80	13 29 54 19 28
58 54 16 24 15	51 54 44 82 00	62 61 85 04 69	38 18 65 18 97	85 72 13 49 21
34 85 27 84 87	61 48 84 56 26	90 18 48 13 26	37 70 15 42 57	65 65 80 39 07
03 82 18 27 46	57 99 16 96 56	30 33 72 85 22	84 64 38 56 98	99 01 30 98 64
62 95 30 27 59	37 75 41 66 48	86 87 80 61 45	23 53 04 01 63	45 76 08 64 27
08 45 93 15 22	60 21 75 46 91	98 77 27 85 42	28 88 61 08 84	69 62 03 42 73
70 08 55 18 40	45 44 75 13 90	24 84 96 61 02	57 55 66 83 15	73 42 37 11 61
01 85 89 95 86	51 10 19 34 88	15 84 97 18 75	12 76 39 43 78	64 63 91 08 25
72 84 71 14 35	19 11 58 49 26	50 11 17 17 76	86 31 57 20 18	95 60 78 46 75
88 78 28 16 84	13 52 53 94 53	75 45 69 30 96	73 89 65 70 31	99 17 43 48 76
45 17 75 65 57	28 40 19 72 12	25 12 74 75 67	60 40 60 81 19	24 62 01 61 16
96 76 28 12 54	22 01 11 94 25	71 96 16 16 88	68 64 36 74 45	19 59 50 88 92
43 31 67 72 30	24 02 84 08 63	38 32 36 66 02	69 36 38 25 39	48 03 45 15 22
50 44 86 44 21	66 06 58 05 62	68 15 54 35 02	42 35 48 96 32	14 52 41 52 48
22 66 22 15 86	26 63 75 41 89	58 42 36 72 24	58 37 52 18 51	03 37 18 39 11
96 24 40 14 51	23 22 30 88 57	95 67 47 29 83	94 69 40 06 07	18 16 36 78 86
31 73 81 61 19	60 20 72 83 48	98 57 07 23 69	65 95 39 69 58	56 80 30 19 44
78 60 73 89 84	43 89 84 36 45	56 69 47 07 41	90 22 91 07 12	78 35 34 08 72
84 37 80 61 56	70 10 23 98 05	85 11 34 76 60	76 48 45 34 60	01 64 18 39 96
36 67 10 08 23	98 93 35 08 86	89 28 76 29 81	33 34 91 58 93	63 14 52 32 52
07 28 59 07 48	89 64 58 89 75	83 85 62 27 89	30 14 78 56 27	86 63 59 80 02
10 15 83 87 60	78 24 31 66 56	21 48 24 06 93	91 98 94 05 49	01 47 59 38 00
55 19 68 97 65	03 73 52 16 56	00 53 55 90 27	33 42 29 38 87	22 13 88 83 34
53 81 29 13 39	35 01 20 71 34	62 33 74 82 14	53 73 19 09 03	56 54 29 56 93
51 86 32 68 92	33 98 74 66 99	40 14 71 94 58	45 94 19 38 81	14 44 99 81 07
35 81 70 29 13	80 03 54 07 27	86 84 78 32 66	50 85 52 74 33	13 80 55 62 54
37 71 67 95 13	20 02 44 95 84	64 85 04 05 72	01 32 90 76 14	53 89 74 60 41
93 66 13 83 27	82 78 64 64 72	28 54 86 53 84	48 14 52 88 84	56 07 93 89 30
02 96 08 45 65	13 05 00 41 84	83 07 54 72 59	21 45 57 09 77	19 48 56 27 44
49 83 43 48 35	82 88 33 69 96	72 36 04 19 96	47 45 15 18 60	82 11 08 95 97
84 60 71 62 46	40 80 81 30 37	34 39 23 05 38	25 15 35 71 30	88 12 57 21 77
18 17 30 88 71	44 81 14 88 47	89 23 30 63 15	56 34 20 47 88	89 82 93 24 98
79 69 10 61 78	71 32 76 95 62	87 00 22 58 62	82 54 01 75 26	43 11 71 89 31
75 93 36 57 83	56 20 14 82 11	74 21 97 90 65	86 42 68 63 86	74 54 13 26 84
38 30 92 29 03	06 28 81 39 38	62 25 06 84 63	61 28 08 83 67	04 32 82 08 09
51 28 60 10 34	31 67 75 85 80	51 87 02 74 77	76 15 48 48 44	18 55 63 77 09
21 31 38 86 24	37 78 38 86 24	37 78 81 53 74	73 24 16 10 33	70 47 14 54 36
29 01 23 87 88	58 02 39 37 67	42 10 14 20 92	16 55 23 42 46	64 96 08 11 06

In a copy of the first table (Table H.3), for illustrative purposes, we have numbered the horizontal blocks as 1–5 and those blocks from the top downward as 1–10.

Table H.3 Random Number Table (#1) Designates

	1	2	3	4	5
1	97 74 24 67 62	42 81 14 57 20	42 53 32 37 32	27 07 36 07 51	24 51 79 89 73
	16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59	88 97 54 14 10
	12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21	88 26 49 81 76
	03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 28 16 80 45	60 11 14 10 95
	55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	08 18 44 32 53	23 83 01 30 30
2	16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	87 35 20 96 43	84 26 34 91 64
	84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	21 76 33 50 25	83 92 12 06 16
	63 01 63 78 59	18 95 55 67 19	98 10 50 71 75	12 86 73 58 07	44 39 52 38 79
	33 21 12 34 29	78 64 58 07 82	52 42 07 44 38	15 51 00 13 42	99 66 02 79 54
3	57 60 86 32 44	09 47 27 96 54	49 17 48 09 62	90 52 84 77 27	08 02 73 43 28
	18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	06 76 50 03 10	55 23 64 05 05
	28 62 38 97 75	84 16 07 44 99	83 11 46 32 24	20 14 85 88 45	10 93 72 88 71
	23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	32 98 94 07 72	93 85 79 10 75
	62 38 28 19 95	50 92 26 11 97	00 56 76 31 38	80 22 02 53 53	88 60 42 04 53
4	37 85 94 35 12	83 39 50 08 30	42 34 07 96 88	54 42 06 87 98	35 85 29 48 39
	70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	17 76 37 13 04	17 74 21 19 30
	56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	70 33 24 03 54	97 77 46 44 80
	99 49 57 22 77	88 42 95 45 72	16 64 36 18 00	04 43 18 66 79	94 77 24 21 90
	18 08 15 04 72	33 27 14 34 09	45 59 34 68 49	12 72 07 34 45	99 27 72 95 14
5	31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	52 85 66 60 44	38 68 88 11 80
	68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52	68 07 97 06 57
	74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76	15 54 55 95 52
	27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	96 46 92 42 45	97 60 49 04 91
	00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 66 04 26	11 04 96 67 24
6	29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	34 25 20 57 27	40 48 73 51 92
	18 90 82 66 59	83 62 64 11 12	67 19 00 71 74	60 47 21 29 68	02 02 37 03 31
	11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	76 70 90 30 86	38 45 94 30 38
	35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	16 92 53 56 16	02 75 50 95 98
	38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	40 01 74 91 62	48 51 84 08 32
7	31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	00 52 43 48 85	27 55 26 89 62
	66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	76 83 20 37 90	57 16 00 11 66
	14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
	68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
	20 46 78 73 90	97 51 40 14 02	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
8	64 19 58 97 79	15 06 15 93 20	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
	05 26 93 70 60	22 35 86 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
	70 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 05 77 09 51
	68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
	26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
9	14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43
	17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	44 95 27 36 99	02 96 74 30 83
	90 26 59 21 19	23 52 23 33 12	96 93 02 18 39	07 02 18 36 07	25 99 32 70 23
	41 23 52 55 99	31 04 49 69 98	10 47 48 45 88	13 41 43 89 20	97 17 14 49 17
	60 20 50 81 69	31 99 73 68 68	35 81 33 03 76	24 30 12 48 60	18 89 10 72 34
10	91 25 38 05 90	94 58 28 41 36	45 37 59 03 09	90 35 57 29 12	82 62 54 65 60
	34 50 57 74 37	98 80 33 00 91	09 77 93 19 82	74 94 80 04 04	45 07 31 66 49
	85 22 04 39 43	73 81 53 94 79	33 62 46 86 28	08 31 54 46 31	53 94 13 38 47
	09 79 13 77 48	73 82 97 22 21	05 03 27 24 83	72 89 44 05 60	35 80 39 84 88
	88 75 80 18 14	22 95 75 42 49	39 32 82 22 49	02 48 07 70 37	16 04 81 67 87
	90 98 23 70 00	39 90 03 08 90	55 85 78 38 36	94 37 30 69 32	90 89 00 76 33

The horizontal and vertical numbers of the columns give us locational designations for determining intersecting axes and, thus, a starting point within the table. To determine the point of axial intersection and to get a starting point, you might engage in any one of several activities. The purpose of any of these is to select two digits purely by chance. Here are some suggestions.

Use a telephone directory. Open a telephone directory at random. Take the last two digits of the first number in the first column on either the left-hand or right-hand page. Let the first digit of those chosen be the designating digit of the horizontal column of the table of random numbers; the second digit will, then, indicate the number of the vertical column. The intersection of the two columns will indicate the block where you will begin within the table to select your random numbers, either in an upward or in a downward succession along the vertical column.

Note a vehicle license plate. Step outside and observe the first vehicle that passes. Note the last (or first) two digits of the license plate. The first digit will give you the horizontal column designation; the second digit, the vertical column. Where they intersect is your place of beginning within the table.

Look at a dollar bill. Note the first two digits of the serial number in the lower left- or upper right-hand corner to determine the axial locations.

Check the stock quotations. Take any newspaper and turn to the stock quotation page. Take the first letter of your surname. The first stock which begins with that letter will be your predetermined stock; note its quotation for high and low. Disregard the fractional quotations. Take the first two digits in either the high or low quotation column; or if only one digit appears in each column, take the two digits together.

Having arrived at the digital block location, the next step is to determine the size of the proposed sample. If it is to be less than 100 individuals, we shall select only two-digit numbers; if it is to be less than 1,000, we shall need three digits to accommodate the sample size.

Let us go back to the total population for a moment and consider the total group from which the sample is to be drawn. It will be necessary to designate these individuals in some specific manner. It is, therefore, advantageous to arrange the individuals within the population in some systematic order (alphabetically, for example, by surname) and assign to each a serial number for identification purposes.

Now we are ready for the random selection. We shall start with the upper left-hand digits in the designated block and work first downward in the column; if there are not enough digits for the total sample demand in that direction, we will return to the starting digits and proceed upward. Having exhausted all 50 digits in any one column, move to the adjoining columns and proceed as before until the sample requirement is filled. As each digit designation comes up, select the individual in the population who has been assigned that random number. Keep so selecting until the entire sample total is reached.

The following illustration recapitulates what we have been describing. We have pulled a dollar bill from our wallet and note that the first two digits are 4 and 5. These we shall use in locating the beginning block from which we will select the first random number for the sample. For purposes of illustration, we shall assume that the total population consists of 90 individuals from which we shall select a sample group of 40. We will need random numbers of two digits each.

Here again is a reduction of the first table of random numbers (Table H.4), from which we have extracted the beginning block of digits that is at the intersection of the fourth horizontal block column and the fifth vertical block column.

Beginning now in the upper left-hand corner of the designated block, and remembering that there are only 90 individuals in the total population, we see that by going down the two leftmost columns of digits in the block we will choose from the total population individual number 4, and individual number 13. The next two digits (96) do not apply, because there are only 90 individuals in the population, and so we will ignore this number. Our next choices will be 10, 34, 60, 76, 16, 40, and again we will ignore the 00 as well as the next number, 76 (the first set of digits in the third block), because we already have selected that number.

We have perhaps said enough with respect to the use of a table of random numbers, but because randomization is so often effected by the use of just such a table, the foregoing discussion is probably justified.

Table H.4 Random Number Table (#1) with Designated Block

1	2	6	4	5
1	97 74 24 67 62	42 81 14 57 20	42 53 32 37 32	27 07 36 07 51
	16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59
	12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21
	03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 28 16 80 45
2	55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	60 11 14 10 95
	16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	23 83 01 30 30
	84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	84 26 34 91 64
	63 01 63 78 59	18 95 55 67 19	98 10 50 71 75	88 26 49 81 76
3	33 21 12 34 29	78 64 58 07 82	52 42 07 44 38	60 11 14 10 95
	57 60 86 32 44	09 47 27 96 54	49 17 48 09 62	23 83 01 30 30
	18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	84 26 34 91 64
	28 62 38 97 75	84 16 07 44 99	83 11 46 32 24	99 66 02 79 54
4	23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	08 02 73 43 28
	62 38 28 19 95	50 92 26 11 97	00 56 76 31 38	10 93 72 06 16
	37 85 94 35 12	83 39 50 08 30	42 34 07 96 88	12 72 07 34 45
	70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	99 27 72 95 14
5	56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	35 85 29 48 39
	99 49 57 22 77	88 42 95 45 72	16 64 36 18 00	44 39 52 38 79
	18 08 15 04 72	33 27 14 34 09	45 59 34 68 49	55 23 64 05 05
	31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	88 77 46 44 80
6	68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52
	74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76
	27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	97 60 49 04 91
	00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 66 04 26
7	29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	11 04 96 67 24
	18 90 82 66 59	83 62 64 11 12	67 19 00 71 74	40 48 73 51 92
	11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	68 07 97 06 57
	35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	38 45 94 30 38
8	38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	02 75 50 95 98
	31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	48 51 84 08 32
	66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	27 55 26 89 62
	14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	02 02 37 03 31
9	68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	22 98 12 22 08
	20 46 78 73 90	97 51 40 14 02	44 95 92 63 16	07 52 74 95 80
	64 19 58 97 79	15 06 15 93 20	39 54 16 49 36	49 37 38 44 59
	05 26 93 70 60	22 35 86 15 13	01 90 10 75 06	47 95 93 13 30
10	70 97 10 88 23	09 98 42 99 64	40 78 78 89 62	02 67 74 17 33
	68 71 86 85 85	54 87 66 47 54	59 56 78 06 83	58 05 77 09 51
	26 99 61 65 53	58 37 78 80 70	52 27 36 22 41	52 91 05 70 74
	14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	29 56 24 29 48
11	17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	94 44 67 16 94
	90 26 59 21 19	23 52 23 33 12	44 95 27 36 99	82 62 54 65 60
	41 23 52 55 99	31 04 49 69 98	07 02 18 36 07	18 89 10 72 34
	60 20 50 81 69	31 99 73 68 68	24 30 12 48 60	97 17 14 49 17
12	91 25 38 05 90	94 58 28 41 36	35 81 33 03 76	35 80 39 84 88
	34 50 57 74 37	98 80 33 00 91	45 37 59 03 09	45 07 31 66 49
	85 22 04 39 43	73 81 53 94 79	09 77 93 19 82	53 94 13 38 47
	09 79 13 77 48	73 82 97 22 21	33 62 46 86 28	39 32 82 22 49
13	88 75 80 18 14	22 95 75 42 49	05 03 27 24 83	02 48 07 70 37
	90 98 23 70 00	39 90 03 08 90	39 85 78 38 36	16 04 81 67 87
14	34 50 57 74 37	98 80 33 00 91	55 85 78 38 36	90 89 00 76 33
	85 22 04 39 43	73 81 53 94 79	59 85 78 38 36	94 37 30 69 32

Appendix I: Calculating the Standard Deviation

The following are formulas for calculating the standard deviation, although spreadsheet software will calculate a standard deviation for you. Although it looks longer, the formula on the right is considered a shortcut!

$$s^2 = \frac{\sum(y - \bar{y})^2}{n-1}$$

$$\sum(y - \bar{y})^2 = \sum y^2 - \frac{(\sum y)^2}{n}$$

Assume that the average (mean) number of prior arrests of a sample of the jail population has been calculated. The standard deviation is needed to identify how much variation exists in the number of priors. Also assume that there are only 5 cases in the sample (although, in reality, a much bigger sample would be necessary).

Statisticians identify variables as "X" or "Y." They use "X" for variables they are trying to explain (called dependent variables) and "Y" for variables (called independent variables) that they think may cause or have some sort of relationship with "X." In this case, Prior Arrests is a "Y." That means that it is believed that prior arrests cause something, perhaps longer Length of Stay. To calculate the standard deviation of "Y" (Prior Arrests), follow these steps.

Step 1: List each measurement of Y (Prior Arrests).

Table 1.1 Measures of "Y" (Prior Arrests)

Case Number	"Y"
Prisoner #1	5
Prisoner #2	7
Prisoner #3	1
Prisoner #4	2
Prisoner #5	4
Total	19

Step 2: Square each measurement of Y and record it in a second column.

Table I.2 Measures of Y and Y-Squared

Case Number	"Y"	Y ²
Prisoner #1	5	25
Prisoner #2	7	49
Prisoner #3	1	1
Prisoner #4	2	4
Prisoner #5	4	16
Total	19	95

Step 3: Square the total of Y.

$$19 * 19 = 361$$

Step 4: Divide **the** result of Step 3 by the number of cases in the sample.

$$361 / 5 = 72.2$$

Step 5: Subtract the result of Step 4 from the Y² total.

$$95 - 72.2 = 22.8$$

Step 6: Subtract 1 from the number of cases in the sample.

$$5 - 1 = 4$$

Step 7: Divide the result of Step 5 by the result of Step 6.

$$22.8 / 4 = 5.7$$

Step 8: Using a calculator or a square root table, find the square root of the result of Step 7.

$$\sqrt{5.7} = 2.39$$

Step 7 produces the variance, and Step 8 produces the standard deviation

In this analysis, the "average" prisoner has 3.4 prior arrests. However, the standard deviation is 2.39, which suggests there are major differences between the prisoners in the sample. Now, for the shortcut! If the standard deviation is not provided, and if there is not sufficient time to calculate it, the standard deviation can be **estimated** by dividing the range by 4. Even if all the calculations are done, this is a good check on them.

Appendix J: Calculating Chi Square

Let's use Sheriff #2, introduced earlier, to illustrate another example. The jail has experienced an increase in the number of incidents and the sheriff wants to explore the relationship between the number of incidents prisoners commit and the number of prior arrests they have. Sheriff #2 suspects that some more serious offenders are spending longer periods of time in the facility. He calculates a **Chi Square** to discover the nature of the relationship between these two data elements.

Calculating a Chi Square takes some time, but is straightforward. Follow these steps to determine if there is a relationship between membership in a particular category of one data element and membership in another category of a different data element.

Step 1: Construct a contingency table for the two data elements.

Sheriff #2 constructs a contingency table (Table J.1). In this process, the prisoners in the sample are divided into two groups: those who were involved in incidents and those who were not. Then for each inmate, Sheriff #2 determines their number of previous arrests.

Table J.1 Contingency Table of Prior Arrests and Incident Involvement

Prior Arrests	Involved in Incidents	Not Involved in Incidents
More than 5	30	15
3 - 5	25	20
0 - 2	25	65

This is a contingency **table** that shows the observed frequencies.

Step 2: Total the columns and the rows and write the totals around the outside of the contingency table.

Table J.2 Contingency Table of Prior Arrests and Incident Involvement with **Marginals**

Prior Arrests	Involved in Incidents	Not Involved in Incidents	
More than 5	30	15	45
3 - 5	25	20	45
0 - 2	25	65	90
	80	100	180

Lines of numbers are called "columns" if they are vertical and "rows" if they are horizontal. The numbers written around the outside of the contingency tables are called "**marginals**," because they are written in the margins (Table J.2).

Step 3: Calculate the expected distribution as if there were no relationship between the two data elements.

Of the 180 inmates in Sheriff #2's sample, half of the inmates (90) had from 0 to 2 prior arrests, while one quarter of them (45) had from 3 to 5 prior arrests, and a final quarter (45) had more than 5 prior arrests. Of the 180 inmates, 80 were involved in incidents at the jail and 100 were not. If there were no relationship between the number of prior arrests and committing incidents at the jail, the arrest pattern within the two groups based on incident status should be the same as the pattern for all the inmates. The expected distribution then would look something like Table J.3.

Table J.3 Contingency Table of Prior Arrests and Incident Involvement with Expected Distribution

Prior Arrests	Involved in Incidents	Not Involved in Incidents	
More than 5	20	25	45 = A
3 - 5	20	25	45 = B
0 - 2	40	50	90 = C

Table J.3 shows the Expected Frequencies. Sometimes, the **patterns** are not quite so clear as in this example. When that is the case, a formula can be used to calculate the expected frequency of each "box" in the contingency table. Statisticians call each "box" a "cell."

Expected frequencies of the group who were involved in incidents would be:

$$\begin{aligned}M * (A / T) \text{ or } 80 * 45 / 180 &= 80 * .25 = 20 \\M * (B / T) \text{ or } 80 * 45 / 180 &= 80 * .25 = 20 \\M * (C / T) \text{ or } 80 * 90 / 180 &= 80 * .25 = 40\end{aligned}$$

Expected frequencies of the group who were **not** involved in incidents would be:

$$\begin{aligned}M * (A / T) \text{ or } 100 * 45 / 180 &= 100 * .25 = 25 \\M * (B / T) \text{ or } 100 * 45 / 180 &= 100 * .25 = 25 \\M * (C / T) \text{ or } 100 * 90 / 180 &= 100 * .25 = 50\end{aligned}$$

Step 4: Assign a number to each cell in the contingency table.

When this is done, the contingency table will look something like Table J.4.

Table J.4 Contingency Table with Assigned Cell Numbers

Prior Arrests	Involved in Incidents	Not Involved in Incidents	
More than 5	Cell 1 20	Cell 4 25	45 = A
3 - 5	Cell 2 20	Cell 5 25	45 = B
0 - 2	Cell 3 40	Cell 6 50	90 = C

80 = M 100 = N 80 = T

This step is to make sure that you identify **each** cell and avoid confusion.

Step 5: For each cell, compute the formula.

$$\frac{(O - E)^2}{E}$$

(Observed Frequency - Expected Frequency)² / Expected Frequency

So, for Cell #1, after using the formula, the results are:

$$\begin{aligned}(30 - 20)^2 / 20 \\ \text{or} \\ (10)^2 / 20 \\ \text{or}\end{aligned}$$

$$100 / 20 = 5$$

Taylor Fitz-Gibbon and Moms suggest making a worksheet like Table J.5 for each of the items in the formula to make sure that nothing is overlooked.

Table J.5 Contingency Table Worksheet

1 Cell	2 O	3 E	4 O - E	5 O - E²	6 O - E² / E
1	30	20	10	100	5.000
2	25	20	5	25	1.250
3	25	40	-15	225	5.625
4	15	25	-10	100	4.000
5	25	20	5	25	1.250

Step 6: Calculate the total of Column #6.

The total of column #6 is the **Chi Square**.

Step 7: Calculate the Degrees of Freedom.

"Degrees of Freedom" is a mathematical concept that measures the amount of information available in normally distributed data. While this may be hard to explain, it is easy to do. Block out one row and one column of your original contingency table and count the remaining cells. The result looks like Table J.6.

Table J.6 Contingency Table Showing Degrees of Freedom

Labels	Column #1	Column #2
More than 5	①	
3 - 5	②	

In this case, there are only two cells left. Therefore, this contingency table has only **2** degrees of freedom.

Step 8: Consult a Chi Square table.

Statistics texts have tables of Chi Squares that will tell you for a particular significance level what the Chi Square is for contingency tables with the same Degrees of Freedom. At the .05 level, with two degrees of freedom, Chi Square = **5.99**.

Step 9: Determine whether a significant relationship exists between the two variables.

If the Chi Square that you calculated is larger than the Chi Square you found in the table, then there is a statistically significant relationship between the two variables. In our example, because the calculated Chi Square is 21.625 and the table value of Chi Square is 5.99, Sheriff #2 is certain that a significant relationship exists between the number of prior arrests and being involved in incidents while in jail.

Appendix K: Manual Data Display

Bar Charts

Hint 1: Use graph paper.

Graph paper, gridded to 10 x 10, on which the graph is built in units of 10, seems easiest to use for most jail data. Each small square can represent 1, 10, 100, or whatever the most appropriate number is of the items being counted. Graph paper also comes in what is called "non-reproducible blue" so that the graph lines on the paper do not show up when it is copied.

Hint 2: Calculate the height of the tallest bar first.

The height of the tallest "bar" will determine the height of the Y Axis. The Y Axis should be at least a bit taller than the highest bar.

Hint 3: Consider different scales.

Each small square on the graph paper represents a fixed number of cases. The easiest way to do this is to determine how many squares are on the paper from top to bottom or side to side. Then look at the largest category. Make sure that the scale selected will accommodate the bar representing the largest category. The scale that is used often determines how much impact the graph has.

Hint 4: Determine the width of the bars.

Decide both how wide and how far apart the bars will be. Mark them on the X Axis. It is wise to avoid making tall, thin bars because they look like they are going to fall down. If so many categories are included that the width of the paper dictates that the bars be thin, consider turning the paper around so that the Y Axis is longer than the X Axis. If the bars are now too tall for the paper, use another scale, or use larger paper and reduce the finished chart.

Hint 5: Draw each bar in pencil.

Using the scale on the Y Axis, make a pencil line where the top of each bar will be, then draw in the rest of each bar in pencil. Check your work against the scale and the data. If everything appears to be right use a pen or graphic tape to finish the bars.

Hint 6: Label the chart.

Put a title on the chart, label each of the categories, and mark the scale on the Y Axis.

Hint 7: Consider using color or shading the bars.

Either shading or coloring the bars will make the chart easier to interpret.

Pie Charts

Hint 1: Invest in a drafting compass.

Use a compass to make a circle sized properly for your paper. If you have no compass, you can make a pie chart with a glass or other round object.

Hint 2: Use a photocopy.

Make a copy of the circle and work on the copy. That way, when mistakes are made, it is not necessary to start again from the beginning.

Hint 3: Calculate the size of each piece of the pie.

Sometimes, it is easy to calculate the size of the segments. If the category is expressed in a percentage and the values are 50%, 25%, and 25%, it is obvious that half a circle will represent 50%, etc. Often, however, it is not so obvious. When in doubt, since there are 360 degrees in a circle and the total number of cases in each category is also known, it is always possible to resort to algebra and geometry. To figure out how many degrees to include in the piece of pie that represents a particular category, let "X" stand for the number of degrees in the piece of the pie and substitute that number in the following formula:

$$\frac{X}{360} = \frac{n \text{ (category)}}{N \text{ Total}}$$

Do this for each piece of the pie. Use a protractor to mark off the degrees on the circle. Do not just estimate what 51 degrees looks like.

Hint 4: Label the chart.

Label both the whole pie chart and each piece of the pie. Usually labels should be placed around the outside of the pie, with neatly drawn arrows leading to the section the label represents. Another

approach is to develop a key beneath the pie chart to show what each shading or color on the pie represents.

Line Graphs

Hint 1: Use graph paper.

Once again, 10 x 10 graph paper is easiest to use for most jail data.

Hint 2: Draw in the X and Y axes.

Both the axes represent a data element in a line graph.

Hint 3: Determine the range of values to be displayed.

Identify the highest and lowest value of the data element to be displayed on the Y Axis. This determines how tall the Y Axis has to be, since it should be at least a bit taller than the highest value.

Hint 4: Repeat the process for the X axis.

For much jail data, all that is required is to divide the X Axis into equal sections to represent the number of years, months, days, etc., to be included on the graph. If there are actually two data elements, like height and weight, a scale also will be needed for the X Axis. Data like this is quite rare in jails, however.

Hint 5: Determine the scale.

This process is identical to developing a scale for a bar chart.

Hint 6: Plot each data point.

Using the scale on the Y Axis, make a dot for each data point to be displayed on the graph. This is called "plotting the data." Be sure to use a pencil and a ruler to place the points in the right spots. The farther from the Y Axis, the easier it is to make mistakes. Once the data points are plotted, connect the dots in pencil. Check your work against the scale and the data. If there are no errors, use a pen or graphic tape to finish the graph.

Hint 7: Title the line graph.

Put a title on the line graph, label each of the categories, and mark the scale on both the X and Y Axes.

Graphic Tools

A first stop for purchasing drafting equipment and graphics **supplies** should be a local store that sells office **and/or art** supplies and equipment. If there is not one nearby, this material can be ordered from a catalog. Following are some basic supplies and equipment that make working on graphics much easier. While some sheriffs and jail administrators may consider these items a substantial investment, many are quite inexpensive, particularly when compared to the cost of professionally done graphics.

Basic 1: A Place to Work

A good place to work is essential, and while any flat surface like a desk or table will suffice, all drafting students know the value of a drafting table. Many graphic artists work on a "light board," a box with a translucent glass top that is slanted at about 25 degrees. Inside the box is a fluorescent light. The light illuminates anything (like a grid or lined paper) under the **paper** being worked on. This makes it much easier to line things up. Commercial light boards are rather expensive, but they are quite easy and relatively inexpensive to build.

Basic 2: Good Writing Equipment

There are many writing tools on the market. Most expensive—and most effective—are technical or drafting pens, which come in different sizes and make lines of different widths. They can cost more than \$10. Almost as good as drafting pens are fine-point markers, and they are much less expensive. Do not use a ballpoint pen as it does not copy well and tends to leave little "blobs" of ink. Several good possibilities exist for making lines that do not show on the final copy. The best is a non-reproducible blue pencil, priced about the same as a regular colored pencil. However, it will leave an obvious mark on the original copy.

Basic 3: A Good Straightedge

Do not use the edge of a yellow pad. If a drafting table is available, a T square is the best alternative, but a good ruler with a metal indented edge is equally effective. The indented edge prevents smearing when the **ruler** is picked up. Drafting tape is another excellent alternative. It comes in a variety of widths (from 1/64 inch to ½ inch in width) and colors and makes sharp lines. Dotted lines, dashes, and other variations are also available. Drafting tapes vary in cost by width and will last quite some time.

Basic 4: A Good Compass and Protractor

A good compass and protractor are even more essential than a good straightedge for doing pie charts. Another alternative would be to use a **template** (a sheet of plastic or metal with **geometric shapes** punched out) to trace circles, but these are not big enough for most pie charts.

Basic 5: The Right Kind of Paper

There **are** differences in paper. White, relatively smooth paper is preferable because it copies better and is easier to write on. Two special kinds of paper are worth mentioning here:

1. Non-reproducible blue graph paper has lines that do not show when photocopied (however, it may be necessary to lighten up the copy a bit).
2. Drafting paper makes smearing ink difficult (although not impossible).

Basic 6: Something to Make Letters and Symbols

With few exceptions, most people's handwriting is not good enough to go into a document or on a graphic. Print from a good typewriter looks clean and sharp and will work well in most report graphics. However, when projected, typewriter type is not large enough for an audience to read, and the slightly larger or bolder titles and headings that improve the appearance of a graphic cannot be done on a typewriter.

The most reasonable alternative for doing small amounts of graphic work is "rub-on lettering." Over a hundred different styles of rub-on lettering are sold and each style comes in a variety of sizes. It gets its name because it is literally **rubbed off** the sheet it comes on and **on** to another sheet. If you try this, get a small wooden or metal tool called a "burnisher," which gets the letter off the sheet more smoothly than a pencil. One problem with rub-on lettering is that it is difficult and time consuming to evenly space the letters and get them all aligned. Rub-ons also come in a variety of symbols: circles, squares, triangles, arrows, check marks, etc.

Basic 7: Something to Add Color

A wide variety of colored pens and pencils can be used on paper. In addition, special marking pens are made to use on acetate (transparencies). They are certainly the most reasonable in cost, but all fail to provide an "even" color and as a result may detract from the finished product. Color will probably be added to only the most important graphics. The best alternative is "graphic film." This is a thin sheet of colored, translucent material that has adhesive on the back. It can be positioned over something that already has writing on it—for example, a piece of a pie chart—and it will add the color or texture of your choice. Film is made for use on both paper and transparencies.

Basic 8: A Good Adhesive

The days of rubber cement and cellophane tape have gone by. Rubber cement is too messy, and tape makes lines when copied. Mounting adhesive, essentially rubber cement in aerosol form, is a better alternative. It allows repositioning of what has been sprayed.

Appendix L: Automated Case Study Guide

Chapter 7 of this manual consists of two case studies that allow users to practice the art of analysis. In the document, it is structured as an interactive programmed text. The automated version consists of the same case studies, structured in a **WordPerfect** 6.1 macro that allows the user to answer the questions and interact with the computer.

The first case study is located in a file named **CH7CS#1.WPD**; the second case study is found in **CH7CS#2.WPD**. This automated study guide tells you how to use the case studies. Good luck!

Before you start, you should spend a moment to familiarize yourself with the conventions used in both automated files.

- ✓ Keys you will use on your keyboard are indicated in bold typeface. This applies to individual letters and numbers as well as the function keys (e.g., **F1**, **F2**, **F3**).

- ✓ When the cursor is shaped like an hourglass, this indicates that the computer is still working. Wait until the hourglass is gone before attempting to enter any keystrokes.

- ✓ <**Enter**> indicates to press the "Enter" key on your keyboard.

- ✓ **ALT+Page** Down moves you down a page at a time. **ALT+Page** Up moves you up a page at a time.

- ✓ The key combination that you will use during this exercise is **CTRL+F4**. When you see this combination you should hold down the **CTRL** key while pressing the **F4** key. Then release both keys.

- ✓ When you are in the answer portion of this exercise use the Arrow keys $\leftarrow \rightarrow \uparrow \downarrow$ on your keyboard to maneuver in order to read the answer. When you have finished reading the answer press the <Enter> key to take you to the next question.

- ✓ To begin the exercise go to **Question #1**, place your cursor anywhere **within** the question, and hit <Enter>. If you do not have a color monitor go to **Question #1**, place your cursor anywhere **within** the question, and hit <Enter>.

- ✓ When you start the exercise you will be asked to save the document. Choose a name for the document of not more than 8 letters.

- ✓ There are 10 questions in this exercise. If for some reason you need to stop the exercise before you have finished all 10 questions, you may hit Esc to cancel the exercise. The document will automatically be saved with your answers.

- ✓ If an "Error" or "Cancel" condition occurs while you are going through the exercise and you have to exit the exercise document, start again at Question #1 and the program will take you back to where you left off.

- ✓ If an "Error" or "Cancel" condition occurs in which you have to leave your word processing application, you will need to start the exercise again.