US Army Corps of Engineers
Water Resources Support Center
Institute for Water Resources

AD-A271 723

NATIONAL ECONOMIC DEVELOPMENT PROCEDURES MANUAL - PUBLIC SURVEYS

VOLUME I

Use and Adaptation of Office of Management and Budget Approved Survey Questionnaire Items for the Collection of Corps of Engineers Planning Data

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC
ELECTE
NOV 0 1 1993
S B D

20030226207

93-26038

January 1993

IWR Report 93-R-2

93 10 26 067

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE National Economic Development Procedures Manual-Public Surveys Volume I: Use and Adaptation of Office of Management and Budget Approved Survey Questionnaire Items for the Collection of COE Planning Data

5. FUNDING NUMBERS

6. AUTHOR(S)
Mills, Allan S., Stuart A. Davis, Linda K. Peterson and William J. Hansen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Corps of Engineers
Water Resources Support Center
Institute for Water Resources
Casey Building, #2594
Fort Belvoir, Virginia 22060-5586

8. PERFORMING ORGANIZATION REPORT NUMBER
IWR Report 93-R-2

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Headquarters U.S. Army Corps of Engineers
CECW-PD
Washington, DC 20314-1000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Available from Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release: Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The main purpose of this manual is to provide guidance for the use of The U.S. Office of Management and Budget approved Corps of Engineer Compendium of survey questionnaire items for the collection of planning data. It provides specific guidance on cross referencing the Compendium of approved survey questionnaires by topic of study, methods of data collection, and types of survey questions. It also provides general survey implementation and analysis guidance, supplementing earlier manual coverage of the survey research process as it relates to use of the Compendium of survey questionnaires. For this latter purpose the manual may be used with or without the Compendium.

14. SUBJECT TERMS
Public Surveys, Questionnaires, Statistical Analysis, Survey Design, Sampling

15. NUMBER OF PAGES
148

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

NATIONAL ECONOMIC DEVELOPMENT PROCEDURES MANUAL - PUBLIC SURVEYS

Volume I

Use and Adaptation of Office of Management and Budget Approved Survey Questionnaire Items for the Collection of Corps of Engineers Planning Data

by

Allan S. Mills
Stuart A. Davis
Linda K. Peterson
William J.Hansen

U.S. Army Corps of Engineers
Water Resources Support Center
Institute for Water Resources
Ft. Belvoir, Virginia 22060-5586

January, 1993

IWR Report 93-R-2

PREFACE

This manual was prepared as part of the National Economic Development (NED) Procedures Manual Work Unit within the U.S. Army Corps of Engineers (COE) Planning Methodologies Research Program. Mr. William Hansen of the COE Water Resources Support Center (WRSC), Institute for Water Resources (IWR), manages this Work Unit under the general supervision of Mr. Michael Krouse, Chief, Technical Analysis and Research Division; Mr. Kyle Schilling, Director. IWR; and Mr. Kenneth Murdock, Director, WRSC. Mr. Robert Daniel Chief, Economic and Social Analysis Branch (CECW-PD) and Mr. William Hunt, CECW-PD, are the Technical Monitors for Headquarters, COE.

Dr. Allan Mills, School of Community and Public Affairs, Virginia Commonwealth University, was the principal author and editor of this manual while serving under the terms of an Intergovernmental Personnel Act appointment with IWR. Mr. Stuart Davis, Planner, IWR, provided considerable input to Chapter IV (Survey Implementation), and was mainly responsible for the content of Chapters V (Data Editing) and VII (Report Writing). Ms. Linda Peterson, Statistician, COE District, Mobile, was mainly responsible for the content of Chapter VI (Data Analysis). Mr. William Hansen provided critical assistance with the initial design of this manual and with final revisions to its form and content.

The authors are grateful to the following individuals for reviewing the preliminary draft of this manual and providing

valuable comments and suggestions for improvement: Jon Brown (COE, Buffalo District), Bruce Carlson (COE, St. Paul District), James M. Davenport (Virgi[illegible] Commonwealth University), Ronald W. Hodgson (California State University, Chico), Harry Kelejian and Kamala Rajamani (University of Maryland), Dennis Robinson (COE, IWR), and Ed Rossman (COE, Tulsa District).

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Avail and/or
Dist Special
A-1

ITY INSPECTED 5

TABLE OF CONTENTS

PREFACE iii

CHAPTER I: INTRODUCTION 1

BACKGROUND 1
PURPOSE 2
SCOPE 2
INTENDED AUDIENCE 3
SURVEY IMPLEMENTATION REQUIREMENTS 4
THE PAPERWORK REDUCTION ACT 4
OMB APPROVAL PROCESS 4
DIVISION APPROVAL 5
THE PRIVACY ACT 6
STEPS IN SURVEY PROCESS 7
ORGANIZATION OF THE MANUAL 11

CHAPTER II: CROSS-REFERENCING THE COMPENDIUM 13

COMPENDIUM SURVEY TOPICS 13
CROSS-REFERENCING THE COMPENDIUM 14
TOPICAL SECTIONS 14
SIMILAR QUESTIONS FROM DIFFERENT TOPICAL SECTIONS . 17
TYPES OF SURVEY QUESTIONNAIRES 18
TYPES OF QUESTIONS 19

CHAPTER III: ADAPTING AND REVISING COMPENDIUM QUESTIONS . . 23

ADAPTING TO A MAIL FORMAT 23
ADAPTING TO A TELEPHONE FORMAT 26
REVISING QUESTION WORDING AND RESPONSE FORMATS 27
VALIDITY AND RELIABILITY 27
NEED FOR HIGHER LEVEL MEASUREMENT 30
NEW QUESTIONS FOR APPROVED TOPICAL AREAS 32

CHAPTER IV: SURVEY SAMPLING AND IMPLEMENTATION 37

SAMPLING PROCEDURES 38
DEFINING THE SAMPLE FRAME 39
DETERMINING SAMPLE SIZE 40
CHOOSING THE SAMPLING METHOD 43
PRETESTING 48
PRETESTING THE QUESTIONNAIRE 49
PRETESTING THE SAMPLE LISTING 49
PRETESTING THE METHOD OF SURVEY ADMINISTRATION . . . 51
PERSONNEL MANAGEMENT 51
FACE-TO-FACE SURVEYS 52
INTERVIEWER SELECTION 53
INTERVIEWER TRAINING 54
MONITORING INTERVIEWERS 55

Table of Contents (Continued)

TELEPHONE SURVEYS 55
INTERVIEWER SELECTION, TRAINING, AND MONITORING . . 56
CATI SYSTEMS 57
MAIL SURVEYS 57
EMPLOYEE SELECTION AND TRAINING 60
PROCEDURAL GUIDE TO IMPLEMENTATION 61
GENERAL INSTRUCTIONS 62
SPECIFIC INSTRUCTIONS 63
CODING INSTRUCTIONS 64

CHAPTER V: DATA EDITING AND ENTRY 65

CLERICAL EDITING 65
DETERMINING IF A SURVEY FORM IS USEABLE (ROUND ONE) 65
CORRECTING RECORDING PROBLEMS 67
DATA ENTRY 67
INTERVIEWER CODING AND DATA ENTRY 67
SCANNING 68
SPOT CHECKS OF CODING 69
COMPUTER EDITING 69
SETTING LIMITS FOR OUTLIERS 69
CHECKING COMPLIANCE WITH FILTER QUESTIONS 70
DETERMINING IF A SURVEY FORM IS USEABLE (ROUND TWO) 71

CHAPTER VI: DATA ANALYSIS 73
PERFORMING STATISTICAL ANALYSIS 74
DESCRIPTIVE PROFILES 74
HYPOTHESIS TESTING 74
LEVELS OF MEASUREMENT 77
NOMINAL-LEVEL MEASUREMENT 77
ORDINAL-LEVEL MEASUREMENT 78
INTERVAL-LEVEL MEASUREMENT 78
RATIO-LEVEL MEASUREMENT 79
LEVEL OF MEASUREMENT AND APPROPRIATE STATISTICAL ANALYSES 79
SELECTION BIAS AND DIFFERENTIAL WEIGHTING 79
MISSING DATA 82
REPORTING MISSING DATA CASES 83
DELETING MISSING DATA CASES 83
ASSIGNING VALUES TO MISSING DATA CASES 84
BASIC STATISTICAL ANALYSIS PROCEDURES 86
DESCRIPTIVE UNIVARIATE STATISTICS 87
BIVARIATE RELATIONAL ANALYSIS 94
MULTIVARIATE MODELING PROCEDURES 106

Table of Contents (Concluded)

CHAPTER VII: REPORT WRITING 121
CRITICAL ITEMS TO REPORT 121
ELEMENTS OF A DETAILED REPORT 122
EXECUTIVE SUMMARY 122
BASIC OUTLINE FOR A DETAILED REPORT 122
INTRODUCTION 123
GENERAL DISCUSSION OF SURVEY METHODOLOGY 123
DESCRIBE DATA ANALYSIS AND REPORT RESULTS 126
DESCRIBE IMPLICATIONS 127
APPENDICES 127

CHAPTER VIII: SUMMARY AND RECOMMENDATIONS 129
SUMMARY 129
RECOMMENDATIONS 130

LITERATURE CITED 133

APPENDIX: GLOSSARY OF COMMON STATISTICAL ANALYSES 135
PARAMETRIC STATISTICAL TERMS 137
NON-PARAMETRIC STATISTICAL TERMS 139

LIST OF TABLES

Table Number | Page Number

1. Compendium Questionnaire Topics by Types of Surveys and by Types of Questions 15

2. Method 1: Reporting Omissions 83

3. Method 2: Deleting Missing Cases 84

4. Method 3: Contingent Probability Estimation 85

LIST OF FIGURES

Figure Number | Page Number

1. Univariate Statistics for Survey Data 88

2. Skewed vs. Symmetric Bell-Shaped Distribution 94

3. Bivariate Relational Tests and Statistics for Survey Data 95

4. Elements of Simple Linear Regression 105

5. Multivariate Statistical Proceedings for Survey Data . . 107

CHAPTER I

INTRODUCTION

BACKGROUND

The U.S. Army Corps of Engineers, like other Federal agencies, is required to obtain approval from the Office of Management and Budget (OMB) for Federally funded surveys of ten or more members of the public. The Corps has implemented an approval process with OMB which involves a review, every three years, of the types of survey instruments and data items required for Corps planning surveys. This three year approval process insures that Corps planners can efficiently and effectively implement surveys in a timely manner, without undue delay resulting from the OMB approval requirement.[1]

OMB approved questions are contained in Approved Questionnaire Items for Collection of Planning Data, 1984 (the Compendium). The Compendium consists of previously used survey instruments and questions, grouped into 14 different topical categories. One or more questionnaire types, containing a variety of different data items, are included within each of the topical categories. Corps economists and other planners responsible for conducting surveys can select from the approved questionnaires and, if needed, adapt them for particular types of study objectives.

---

[1] In addition to the three year review and approval process, it is possible to submit additional questions at any time for approval on an ad hoc basis.

PURPOSE

The main purpose of the present manual is to provide guidance for the use of OMB approved Corps survey questionnaires. It provides specific guidance on cross referencing the Compendium of approved survey questionnaires by topic of study, methods of data collection, and types of survey questions. It also provides general survey implementation and analysis guidance, supplementing earlier manual coverage of the survey research process as it relates to use of the Compendium of survey questionnaires. For this latter purpose the manual may be used with or without the Compendium.

SCOPE

This manual builds upon earlier IWR publications in the National Economic Development (NED) Procedures Manual series. These manuals include the NED Urban Flood Damage Procedures Manual - Volume II (Mills et al. 1991), and the NED Recreation Procedures Manuals - Volumes II and III (Moser and Dunning 1986; Hansen et al. 1990). The NED Urban Flood Damage Manual - Volume II describes the design and implementation of surveys to assess residential flood damage, using two Corps flood damage surveys as examples. It includes copies of the questionnaires designed for those surveys and has a detailed description of the first seven steps of the survey process (described below). The NED Recreation Procedures Manuals (Volumes II and III) include examples of survey questionnaires, a detailed discussion of survey sampling (Volume II), and a description of the steps of a contingent valuation survey conducted for planning (Volume III).

The scope of the present manual is thus intended to build on these previous manuals in terms of the steps of the survey process. It first guides the user in hands-on use of the Compendium of approved survey questionnaires. Particular emphasis is placed upon cross-referencing the Compendium to find appropriate types of desired questionnaires and survey items, and on adapting and revising these questionnaires and items to specific survey needs. The manual then addresses survey data analysis and the reporting of survey results, topics not covered in detail in the previous manuals. However, the manual is not meant to be an exhaustive treatment of these topics. Some familiarity with the survey method by the reader is assumed, and basic statistics and survey texts or previous manuals should be consulted for more depth on particular topics.

INTENDED AUDIENCE

This manual is intended as a guide for Corps planners, managers and others who use the Corps Compendium of OMB approved questionnaires. This may include non-Federal sponsors in Corps studies who must design and implement planning surveys. Non-Corps planners and managers who do not have the Corps Compendium may also find this manual useful, particularly the latter chapters concerning analyzing and reporting survey results.

SURVEY IMPLEMENTATION REQUIREMENTS

THE PAPERWORK REDUCTION ACT

The requirement for obtaining OMB approval for surveys is found in Public Law 96-511, the Paperwork Reduction Act of 1980. The law mandates that any survey in which the Federal government solicits response from more than ten individuals, businesses, or organizations outside the government, can only be conducted using OMB-approved questions. A user may extract any combination of OMB-approved questions from the Compendium for the problem being evaluated. The questions can be reworded and rearranged as needed to appropriately reflect the specific planning context or the method of questionnaire administration. Wording changes can also be made to make the questions clearer to the respondent, as long as the intent of the changed wording is the same as the original intent of the questions.

The Paperwork Reduction Act provisions are intended to limit governmental intrusion into the time and personal lives of the general public. Prior to its enactment, there was particular concern that the burden of government paperwork might constitute a significant cost to small businesses. Questions now allowed may be used only to fulfill a specific planning purpose as mandated by laws and regulations.

OMB APPROVAL PROCESS

The Compendium of questionnaire items is the package the Corps of Engineers uses to obtain blanket OMB approval every three years for its planning study survey questions. Approval can also be obtained for specific individual surveys during that three year cycle.

The Compendium is a notebook of questionnaire items prepared by the Institute for Water Resources. It includes a completed copy of OMB Form 83, which documents the expected number of burden hours on the public of completing questionnaires and the approximate costs to the Government. An accompanying supporting statement documents why obtaining survey data is necessary, how it is to be used, what efforts are required to avoid duplication and minimize the burden hours on the public - particularly small businesses, how the confidentiality of response is to be insured, whether or not sensitive questions are to be asked, and what methods of statistical analysis are to be employed.

After OMB approval is granted for the Compendium of survey items, the OMB approval number and expiration date of the approval must accompany every survey form used. Further details on OMB requirements can be found in 5 CFR 1320 (Section 5 of the Code of Federal Regulations).

DIVISION APPROVAL

After questionnaire items are approved by OMB, their use in individual Corps planning studies still requires internal approval by Corps Division offices. General requirements for Corps survey implementation are contained in Engineering Regulation (ER) 1105-2-100 which states that, "Any particular set of questions to be asked of 10 or more respondents shall be approved by the Division Commander." The following five items are to be included in requests to the Division Commander for survey approval:

(1) The research questions to be answered.
(2) The sampling strategy being employed.
(3) Data collection procedures being employed, and follow up procedures.
(4) Data analysis plan.
(5) Copy of proposed questionnaire.

THE PRIVACY ACT

Implicit in OMB approval and Corps regulation guidelines for conducting questionnaire surveys is compliance with the provisions of the Privacy Act of 1974 (PL 93-579). Those planning to do survey research projects should plan their projects to comply with Privacy Act provisions, including informing respondents: 1) that their participation in a questionnaire survey is completely voluntary; 2) that they have the right to refuse to answer any or all questions; and 3) that all responses are confidential and will not be released outside the agency. They must also be informed that each of their answers will be combined with the responses of others, and that response to questions will only be reported in aggregated form.

A key provision of the Privacy Act is the prohibition against maintaining undisclosed records of information on individuals, businesses, and organizations. Sometimes the nature of a survey makes it necessary to attach and maintain names of respondents with completed survey questionnaires. Normally respondent confidentiality is protected by not associating names with completed questionnaires. Maintaining such information would constitute a "system of records" identifiable to individual respondents. This is only permissible if approval is granted for proper maintenance of the records. To gain this approval a "systems of records notice" must be approved through Department of Defense channels and published in the Federal Register. The approval channels include: 1) the Corps Division office; 2) the information management office at Corps of Engineers Headquarters; 3) the Army Information Systems Command at Fort Huachuca, Arizona; and, 4) the Defense Privacy Office in Arlington, Virginia. Approximately six months should be allowed for this approval process; however, blanket approval can be obtained for a recurring use of the same survey.

Defense Privacy Office requirements (p. 11-1) state that the actual notice of the "system of records" should include the following:

1) Name and location of the new system;
2) The categories of individuals on whom records are to be kept;
3) The categories of records maintained in the system;
4) The routine use of the records contained in the system;
5) The agency's policies and practices for storage, retrievability, access controls, retention, and disposal of records;
6) The title and address of the system manager;
7) The agency procedures for notifying the individuals about whom records are maintained;
8) The agency procedures for granting access and amendment to the records;
9) The categories of sources in the records system.

STEPS IN SURVEY PROCESS

Before attempting to use or adapt questionnaires approved by OMB and included in the Compendium, it is essential to understand some general considerations that apply to all surveys. The basic steps of survey research will therefore be reviewed for the reader before discussing how to use the Compendium of approved questionnaires. The following steps are included in the survey research process:

1) Define Objectives.
2) Select Survey Method.
3) Design Questionnaire.
4) Pretest Questionnaire and Procedures.
5) Draw the Sample.
6) Select and Train Personnel.
7) Collect Data.
8) Assess Non-Response.
9) Code, Enter, and Edit Data.
10) Analyze Data.
11) Write Final Report.

Some of these steps differ for different methods of conducting the survey: face-to-face, mail, telephone, or some combination. The order of this step sequence is also sometimes varied, and in some surveys several of the steps may occur simultaneously. Each of the steps is described below.

Step 1 - Define Objectives: Defining objectives should be the first step in every survey. In order to do this effectively, the analyst must have a good understanding of the basic study problem or purpose for which data must be acquired. Each survey objective should be a specification of one or more types of data or information which the survey will be designed to provide. The Volume II NED Urban Flood Damage Procedures Manual (1991) provides guidance on this step.

Step 2 - Select Survey Method: The principal alternative survey methods are face-to-face, telephone, and mail surveys. The Volume II NED Recreation Manual (1986) provides a comparative discussion of the strengths and weaknesses of each of these methods.

Step 3 - Design Questionnaire: Designing the survey questionnaire involves selecting and formatting the appropriate types of questions necessary for acquiring desired survey information. A questionnaire must be designed differently for different methods of survey administration. The Volume III NED Recreation Manual (1990) and the Volume II NED Urban Flood Damage manual (1991) both provide detailed information on survey design. The task of designing the questionnaire is most complex for mail surveys, where a unified mail-out "package" including a cover letter, mail back materials, and the questionnaire itself must be simultaneously developed. A different kind of complexity is involved in designing telephone questionnaires They must be effective with respondents who only hear the questions and response options read to them.

Step 4 - Pretest Questionnaire and Procedures: Pretesting must be conducted to assess the effectiveness of the survey administration procedures as well as the questions themselves. This is most demanding for mail surveys, where the effectiveness of the entire mail out package must be pretested, including the cover letters and mail out procedures.

Step 5 - Draw the Sample: A survey usually requires drawing a sample of respondents from the population being surveyed. This is done because the population is usually too large for everyone to be surveyed. The sample should be randomly drawn, in order to accurately represent the population surveyed within some tolerable range of sampling error. A discussion of sampling is given in Chapter IV. A more detailed discussion is included in the Volume II NED Recreation Manual on Contingent Valuation (1986).

Step 6 - Select and Train Personnel: Personnel must be selected and trained for interviewing and/or other data collection tasks, depending on the survey method employed. This is discussed in Chapter IV. Interviewers must be trained for face-to-face and telephone surveys. Volume II of the NED Urban Flood Damage Procedures Manual (1991) provides descriptions of the selection and training of survey interviewers. For mail surveys interviewers are not needed, but personnel must be selected and trained to administer the assembly and mailing of the mail out package, as well as the return of completed questionnaires.

Step 7 - Collect Data: Data collection means conducting personal interviews for face-to-face and telephone surveys. For mail surveys, it involves preparing a series of cover letters, assembling other mail out materials (e.g. postage paid return envelopes), and implementing the mailing sequence. The Volume III NED Recreation Manual (1990) and the Volume II NED Urban Flood Damage Manual (1991) both provide detailed information on data collection through different types of survey implementation.

Step 8 - Assess Non-Response: After survey implementation, there is usually some non-response, due to refusals by some potential respondents and to other factors. If this non-response is for random reasons, those who do not respond will be no different than the survey respondents and no bias will be reflected in the survey results. This cannot always be assumed. Therefore a non-response check is necessary. Checks for non-response bias are usually done by attempting to contact a sample of non-respondents after the survey is completed. They are asked why they did not respond and some basic demographic information is collected for comparison to the demographics of those who did respond. Some differences identified can be adjusted for by weighting the survey data prior to analysis. An example of this is provided in the Volume III Recreation Manual (1990). Weighting by only demographics may give false assurance of representative responses if non-response is related to other factors.

Step 9 - Code, Enter, and Edit Data: After data have been collected and assessed for non-response bias, they usually must be coded and entered into computer files. This is a critical step, because of the human error which can occur in transferring data from questionnaires to a computer file. Data are normally keyed into a computer file manually, and then either "verified" manually or keyed in a second time to check for errors. The data are then edited to correct all data entry errors found. A detailed written description of the computer file record layout should be prepared so that anyone who subsequently works with the data can tell exactly which responses coded data represent.

Step 10 - Analyze Data: Although this step comes near the end of the survey process, data analysis should also be envisioned during steps one, two, and three. How the data are to be analyzed to address specific objectives affects the level of measurement needed, and, therefore, the format of the response categories. This will be discussed in more detail in Chapter VI.

Step 11 - Write Final Report: The final step of survey research is to document survey results by preparing some type of written summary or final report. The purpose of this step is to communicate the survey results to a targeted audience who will read and make use of it. Graphical displays of results (figures) often communicate better than results presented in tables. Supporting text should always be included to interpret the meaning of the survey results presented in each table and figure. The written summary or final report should also include a methodology section which explains how the survey was conducted, addressing each of the preceding steps of the survey process. It is also advisable to attach a copy of the survey questionnaire.

ORGANIZATION OF THE MANUAL

The following chapters in this manual address several general topical areas that should be of interest to the potential user of the Compendium. Chapter 2 describes how to cross-reference the contents of the Compendium by topic of study, by different methods of survey data collection, and by different types of survey questions. Sometimes the most appropriate question for a given information need can actually be found in one of the Compendium questionnaires located in a seemingly unrelated topical area. The issues of adapting or revising Compendium questions and developing new questions within approved Compendium content areas are addressed in Chapter 3. Chapters 2 and 3 are written to specifically address use of the Compendium.

The remaining chapters also use Compendium examples to illustrate important points, but most of the content of these chapters is general in nature and does not necessarily require Compendium use. Chapter 4 alerts the reader to the need to plan for careful survey implementation and administration, regardless of the method used or questions asked. A cursory discussion of data editing is provided in Chapter 5. Chapter 6 provides a detailed discussion of data analysis. Chapter 7 addresses report writing, and Chapter 8 provides a concluding summary of the entire manual.

CHAPTER II

CROSS-REFERENCING THE COMPENDIUM

COMPENDIUM SURVEY TOPICS

The OMB approved questionnaire items provide examples of survey instruments designed for a variety of general survey topics. A person designing a survey can consult the Compendium and look for questionnaire items or surveys that have previously been used to address topical questions similar to that person's survey objectives. The Compendium is indexed by 14 tabbed topical sections[1], A through N, which are titled as follows:

A. Urban Residential Flood Damage
B. Urban Commercial Flood Damage
C. Agricultural Flood Damage
D. Shore Protection
E. Contingent Value
F. Employment Benefits
G. Water Conservation
H. Waterway Economic Survey
I. Dock and Carrier Survey
J. Shipper Form (Shallow Draft)
K. Terminal Questionnaire
L. Social Impact Assessment
M. Institutional Analysis
N. Small Boat Survey

In most cases the survey questionnaires in the Compendium will not include the ex .t questions needed to address the desired study objectives. Related questions can usually be found, however, which can be modified or used as a generic model for the necessary questions. This is a function of the nature of

---

[1] Three additional topical sections have been submitted for OMB approval in 1992. Pending approval, they will be: Customer Satisfaction (Section O), Recreation Expenditures (Section P), and Environmental Evaluation (Section Q).

survey research, as no two surveys are exactly the same, even when they have the same objectives. Questions from previously used questionnaires must usually be modified somewhat to assure that they will provide valid and reliable survey data. They must also be modified if it is necessary to improve the numerical precision of measurement provided by question responses (discussed in more detail in Chapter III).

CROSS-REFERENCING THE COMPENDIUM

As indicated above, the Compendium is organized into 14 general topical sections. It is generally first referenced by the user choosing the topical area(s) from the table of contents most appropriate for the objectives of the survey being planned. The Compendium can then be cross-referenced in several ways (See Table 1). The possibilities include cross-referencing by the different types of survey questionnaires included in each section (personal interview, telephone, mail), by questions of similar topical content to the topic of interest but found in different topical sections of the Compendium (indicated by letter subscripts in Table 1), and by the different types of survey questions included in each Compendium questionnaire (demographic questions, facts and behavior, ratings and attitudes).

TOPICAL SECTIONS

One or more OMB approved survey instruments are included in each of the fourteen topical sections. The most obvious way to use the Compendium is to simply select and adapt questions from

TABLE 1

Compendium Questionnaire Topics By Types of Surveys and Types of Questions

Compendium Sections:	Type of Survey Included[1]			Types of Questions		
	Face-to-Face	Phone Surveys	Mail Surveys	Demographic	Facts & Behavior	Ratings & Attitudes
A. Urban Residential Flood Damage	YES$_{B,D}$	NO	NO$_{L}$	NO	YES	YES
B. Urban Commercial Flood Damage	YES$_{A,D}$	NO	NO	NO	YES	YES
C. Agricultural Flood Damage	YES	NO	NO	NO	YES	NO
D. Shore Protection	YES$_{A,B}$	NO	NO	NO	YES	YES
E. Contingent Value	YES	NO	YES$_{L,N}$	YES	YES	NO
F. Employment Benefits	YES	NO	NO	YES	YES	NO
G. Water Conservation	YES	NO	YES	YES	NO	YES

[1] Letter subscripts indicate other Compendium topical sections where similar question content for the named Compendium topic may also be found.

Table 1, concluded.

Compendium Sections:	Type Surveys Included[1] Face-to-Face	Phone Surveys	Mail Surveys	Types of Questions Demographic	Facts & Behavior	Ratings & Attitudes
H. Waterway Economic Survey	YES	NO	NO	NO	YES	NO
I. Dock and Carrier Survey	YES	NO	NO	NO	YES	NO
J. Shipper Form	YES	NO	NO	NO	YES	NO
K. Terminal	YES	NO	NO	NO	YES	NO
L. Social Impact Assessment	NO	YES	YES	YES	YES	YES
M. Institutional Analysis	YES	NO	NO	NO	YES	YES
N. Small Boat Survey	NO$_{E}$	NO	YES$_{E,L}$	NO	YES	NO

[1] Letter subscripts indicate other Compendium topical sections where similar question content for the named Compendium topic may also be found.

questionnaires within the topical area of interest. The left hand column of Table 1 lists the 14 different topical areas of the Compendium.

SIMILAR QUESTIONS FROM DIFFERENT TOPICAL SECTIONS

Some topical sections are obviously similar and have some types of questions in common. For example, topical section A, labeled Urban Residential Flood Damage, is similar to topical section B, labeled Urban Commercial Flood Damage. Some questions which could be used in flood damage surveys (Topical Sections A and B) are also found in the Shore Protection questionnaires of Section D and in the Social Impact questionnaires of Section L. A new survey being designed from questions mainly found in topical sections A or B could possibly also adapt some questions from sections D or L. Likewise, new surveys mainly referencing topical sections D or L could possibly adapt some additional questions from sections A or B.

Topical Sections E, L, and N also have some questions which may be of use in more topical areas than the one in which they are found. For example, a Contingent Value survey of recreation facilities which uses the questionnaires or selected questions from section E could also logically include one or more questions from the Small Boat Survey questionnaire in section N. It could also include recreation facilities questions from the Social Impact questionnaire in section L of the Compendium. Similarly, surveys designed using Compendium sections L or N may also include one or more relevant questions from section E.

Compendium items may also be appropriately adapted to topics other than those specifically indicated by the topical section names. For example, "travel cost" studies are conducted to quantify the same kinds of recreational benefits that contingent value studies are used for, but they usually require zip code or address data from survey respondents. The first question in Section A, Urban Residential Flood Damage, asks for the address of respondents. This item could be adapted to provide zip code data for travel cost studies.

TYPES OF SURVEY QUESTIONNAIRES

Examples of the three most common types of survey questionnaires are included in the Compendium, though not all types are included in every topical section. Absence of an example of one or more of the three types of questionnaires from a topical section does not necessarily mean the missing type(s) are not appropriate for that particular survey topic. The three types of survey questionnaires are: face-to-face, mail, and telephone questionnaires.

Face-to-Face Questionnaires. A face-to-face questionnaire is one in which questions are administered by interviewers in a face-to-face situation. Table 1 shows that the most common type of survey instrument found in the different topical sections of the Compendium is the face-to-face questionnaire. Twelve of the fourteen Compendium sections contain examples of face-to-face questionnaires.

Mail Questionnaires. A mail questionnaire is one which must be read and completed by the respondents themselves, and then mailed by them to the person or agency conducting the survey. Respondents may receive the questionnaires in various ways (e.g. hand delivered to them or put under their auto windshield wipers), but most often the questionnaires are mailed to respondents, together with a cover letter and a self addressed postage-paid return envelope. Only four of the sections (E,G,L,N) present examples of mail survey questionnaires. The first of several questionnaires included in Section E is a mail questionnaire. The only questionnaire in Section G is a mail instrument. The first three of the four questionnaires in Section L are mail instruments. The only questionnaire in Section N is also designed for a mail survey.

Telephone Questionnaires. A questionnaire is used by interviewers to read questions to respondents in telephone interviews. Only Section L of the Compendium provides an example of a telephone questionnaire. The fourth and final questionnaire in Section L (Social Impact Assessment) is designed for a telephone survey. The reader is also referred to Appendix B of the Volume II Urban Flood Damage manual (1991), for an example of a short telephone script adapted for two flood damage surveys.

TYPES OF QUESTIONS

One way to classify types of survey questions is in terms of the following three designations of information collected: Demographic, Factual/Behavioral, and Ratings/Attitudes.

Demographic Questions. Demographic questions are found in less than half of the Compendium sections. These are usually questions pertaining to the personal characteristics of questionnaire respondents or their place of residence or work. For example, the first question in the first questionnaire of Section A of the Compendium asks, "What is your residential address or block number?" Such items are found in the following five sections of the Compendium: E, F, G, and L.

Factual/Behavioral Questions. Table 1 shows that the most common type of survey question found within all survey instruments in the Compendium is factual/behavioral. These questions usually denote either the existence, purpose, quantity, dollar value, physical description, or knowledge of things; or a factual description of some type of human behavior. Questionnaires in 13 of the 14 Compendium sections have such questions. Only section G does not.

Rating/Attitude Questions. The next most commonly occurring questions found within the Compendium are those measuring a rating or an attitude along some low to high numerical continuum or intensity of feeling. Rating or attitudinal items are present in questionnaires in six of the fourteen Compendium sections (A,B,D,G,L,M).

An example of a typical rating item is found on page eleven of the first questionnaire in Compendium Section A, Urban Residential Flood Damage. It asks respondents to rate the effectiveness of a flood preparedness plan in terms of the following effectiveness continuum:

__Excellent __Good __Fair __Not Effective

Numbers are normally attached to each of the above possible rating choices, either in the questionnaire itself or later for analysis purposes (e.g. 4=Excellent, 3=Good, 2=Fair, and 1=Not Effective).

An example of attitudinal items is found on the last two pages of the last questionnaire of Compendium Section D, Shore Protection. There are twelve items measuring "public preference" for different kinds of possible shore protection actions. Each item asks respondents to indicate the degree to which they favor or oppose a particular action in terms of the following preference continuum:

___Strongly Favor	___Favor Somewhat	___Oppose Somewhat	___Strongly Oppose	___Don't Know

As with the rating responses, numbers are also typically assigned to each part of this attitudinal response continuum (e.g. from 5="Strongly Favor" to 1="Strongly Oppose"). The "Don't Know" response at the end is not part of the preference continuum and would be assigned a unique number (e.g. -1 or 9) so that such responses could be omitted during data analysis. If included in the analysis, "Don't Know" responses will bias the results.

CHAPTER III

ADAPTING AND REVISING COMPENDIUM QUESTIONS

The "administration" of a survey refers to the actual process of implementing it. When the decision is made to administer all or part of a selected Compendium questionnaire to respondents in a manner different from that for which it was designed, changes will usually be necessary. The necessary questionnaire changes will be dictated by the method of administration to be used. Adapting a face-to-face or telephone questionnaire to a mail questionnaire generally requires the most revision. Fewer changes are usually necessary when adapting a face-to-face questionnaire to a telephone questionnaire, or vice versa. This is because an interviewer generally administers the questionnaire for both face-to-face and telephone surveys. In mail surveys the only contact those conducting the survey have with the respondent(s) is written contact, the printed mail questionnaire and a cover letter composed to accompany it. Extra effort and creativity are necessary to ensure that the written message is visually appealing and that all questions are easy to read, understand, and accurately answer.

ADAPTING TO A MAIL FORMAT

For mail surveys, face-to-face or telephone questionnaire items must be revised so that they can be easily read and understood by respondents. For example, most of the questionnaire items in Section A of the Compendium which address urban residential flood damage appear to be written for face-to-face survey administration. One of these Compendium items asks about flood insurance and reads as follows:

Do you have flood insurance?
(NO ANSWER CATEGORIES PROVIDED)

A question like this with no answer categories provided is termed "open-ended", as opposed to a "closed-ended" or structured response question for which respondents are given a choice of two or more answers. For the above question an interviewer could easily record the open-ended answer received, usually a "yes" or "no." An interviewer can also be trained to be prepared to clarify the question for any respondents who asks for clarification, or to record longer responses verbatim.

Questions which ask for this kind of "open" response illustrated above should generally be avoided in questionnaires sent by mail. They require considerable effort from some respondents to decide how to answer and to write down enough words or sentences to adequately express their answer. This encourages a higher non-response rate than would otherwise occur with mail questionnaires. They also often add confusion to the process of data coding and data entry, because they tend to be more difficult and time consuming to code and interpret.

An open-ended question such as the one above can be converted to a closed-ended question, the preferred question type for mail questionnaires, by preparing a clear and exhaustive set of answer categories. These answer categories should be formatted in such a way that there will be no doubt in the respondent's mind about exactly how to answer the question (i.e. no need for clarification). Below is how the above item could be reformatted for a mail survey on urban residential flood damage:

Do you <u>now</u> have flood insurance for this residence:

- On the Building(s)? (Circle One Number)

  1. NO
  2. YES
  3. DON'T KNOW

- On the Contents? (Circle One Number)

  1. NO
  2. YES
  3. DON'T KNOW

In this revision two question marks are used, one at the end of each of the two components of the question to make it clear that two responses are necessary. The words "Circle One Number" are put in parentheses at the end of each of these two parts of the question, to make it perfectly clear how respondents are expected to express their answers. "YES" and "NO" answer categories are provided for each of the two possible types of flood insurance -- insurance on residential building(s) and insurance on the contents of the buildings. A "DON'T KNOW' response is also provided to exhaust all possible response options and to discourage guessing by respondents unsure of which answer is correct. Responses are numbered because single digit numbers are easier for respondents to circle than longer words. The numbers can also be used as the response codes for data analysis, taking care during analysis procedures to omit the "Don't Know" responses which would be coded "3". The words "<u>now</u>"

and "this residence" are added to indicate that the respondent should use the present time as their frame of reference for answering the question. "Now" is also underlined to draw particular attention to the present time.

ADAPTING TO A TELEPHONE FORMAT

Good telephone interview design involves preparing a "script" of questions which are asked in a conversational manner. Alternative transitional phrases and alternative questions are used to do this, depending upon the respondent's answers to preliminary questions. The above item might be reformatted for a telephone survey script as follows:

A. Do you now have any flood insurance for your residence; By that I mean the buildings or their contents?

1 YES (Ask part "B" below.)
2 NO (Skip part "B". Go To next question.)
3 DON'T KNOW (Skip part "B". Go to next question.)

B. Does that insurance cover your buildings, the contents of your buildings, or both?

The Building(s)?	(Circle)	1. YES	2. NO	3. DK
The Contents?	(Circle)	1. YES	2. NO	3. DK

Note that the words in parentheses after the responses to Part A above are instructions for the interviewer, telling when to ask and when not to ask the two questions in part B. This saves telephone time and helps maintain respondent interest by making the questioning flow in a conversational manner. This kind of "skip" can be easily programmed when computer-assisted telephone interviewing software is used.

REVISING QUESTION WORDING AND RESPONSE FORMATS

VALIDITY AND RELIABILITY

All survey questionnaire items from the Compendium should be adapted so as to maintain reliability and validity. The way survey questions are worded and structured can influence the reliability and validity of the responses obtained by asking them. Reliable survey questions are those which always produce the same answers from the same respondents when answering under similar circumstances. Valid survey questions are those which always impartially produce the kinds of answers which they are designed to produce; they are questions which measure what they are designed to measure. It is possible to have reliable questions which produce invalid results. Validity and reliability with respect to survey questions are discussed in the Volume III NED Recreation Manual (1990) and the Volume II NED Urban Flood Damage Manual (1991). It is also possible to have a valid and reliable questionnaire, but invalid survey results. This usually happens when the wrong people are surveyed, the sampling is faulty, non-response is high, or when the data collected do not address the survey objectives.

<u>Question Validity.</u> Assuring question validity in survey research is a matter of judgement. The analyst must be constantly on guard against threats to validity. Survey items in the Compendium which appear to be potentially invalid for use in a particular survey must be reworded and/or restructured to improve their validity.

One of the most common causes of invalid questionnaire items is confusion about the unit of analysis for the study. If the unit of analysis is residences or structures, as is common for flood damage surveys (see sections A and B of the Compendium), then care should be taken to ask all questions in such a way that they will tell the analyst something about residences or structures - either directly or indirectly. For example, the first question in the topical section A of the Compendium on urban residential flood damage is, "What is your residential address or block number?" This could pose a threat to validity for absentee property owners included in a mail survey of flooded residential property. The survey could be mailed to "owner" at the address of a flooded residence, but be forwarded to that owner at another location. In such case the respondent's "residential address" would be an invalid response to the intent of this question, verifying the addresses of the flooded residential properties. This threat to validity could be eliminated or reduced by rewording the question to make it clear that the address of the flooded residence is what is desired.

Another common cause of invalid questionnaire items is a biased response format. For example, a question can be easily biased by not exhaustively including all possible response categories. If in doubt, the analyst should provide a final "OTHER" response category for this type of closed-ended question. The following marital status item from Section E of the Compendium questionnaires is a good example:

What is your current marital status? CIRCLE NUMBER

MARRIED......01
SINGLE.......02

The only response alternatives given are "MARRIED" and "SINGLE". Adding responses of "SEPARATED", "DIVORCED", "WIDOWED", and "OTHER" (or simply "OTHER" if the detail is unnecessary) would improve the validity of this item.

Reliability. The most common cause of unreliable survey questions is poor choice of words. Ambiguous words or words with many different meanings cause unreliable results. This is because different respondents answer these questions with different meanings of the word in mind. The analyst can guard against unreliable items by carefully evaluating every word in survey questions selected from the Compendium. Other words should be substituted for any word found to be ambiguous in the context of a question and the items surrounding it. In addition to the analyst's own evaluation, ambiguous question wording is also identified by pretesting; having colleagues and potential respondents answer the questions in a draft questionnaire and asking them afterwards if any wording used was ambiguous to them.

One example of a potentially unreliable question from section L of the Compendium is stated as follows:

How long have you lived here? _______ Years

The word "here" could be interpreted by some respondents to mean the house in which they are living. Others may interpret it to mean the community in which they are living. Still others could interpret it to mean some particular part of the community (e.g. center city or suburb). The possibility of multiple interpretations of this word is potential cause for unreliable question results. If the analyst uses this question to determine how long respondents have lived in their present house, the words in "your present house" should be substituted for "here" to improve reliability.

NEED FOR HIGHER LEVEL MEASUREMENT

It is often the case that a "higher" level of measurement is needed for the planned data analysis than can be obtained from a particular question from the Compendium questions. A detailed discussion of the different levels of measurement and which levels are appropriate for particular data analyses techniques is provided in Chapter VI. Briefly, there are four possible levels of measurement which can be obtained with a survey question. The lowest level is termed "nominal" measurement, where a code number representing a response only serves to indicate a unique name for the response. The next level of measurement is termed "ordinal" because the responses can be numerically ordered above and below one another. The next level of measurement is called "interval" measurement, because numbers assigned to responses indicate the number of equal units (intervals) of distance between different responses. The highest level of measurement is the "ratio" level, which requires a true zero point (e.g. zero dollars is a true zero point; zero degrees Fahrenheit is not) of reference for the numerical measure as well as equal intervals between subsequent units.

An example of revising a question and its response format to obtain higher level numerical data than would have been obtained with the original format comes from a recently conducted Corps urban residential flood damage survey. A question was desired which would measure the amount of formal education of principal wage earners residing in flooded residences, at the interval or ratio level of measurement, so that the data could be used in multiple regression analysis. No education question was found in Section A of the Compendium, the topical section on residential flood damage, but the following question was found in Section E:

> What is the last grade of regular school that you completed -- not counting specialized schools like secretarial, art, or trade schools?
>
> NO SCHOOL
> GRADE SCHOOL (1-8)
> SOME HIGH SCHOOL (9-11)
> HIGH SCHOOL GRADUATE (12)
> SOME COLLEGE (13-15)
> COLLEGE GRADUATE (16)
> POST GRADUATE (17)

The above seven question response categories would produce ordinal level survey data which could be numerically coded using the numbers zero through six. Each higher response code number represents more education than the last, but not in equal amounts because the number of years for each are different.

This question was re-worded as follows to refer to the principal wage earner of the household:

> Please circle the number below which indicates the total years of schooling that the principal wage earner of this household completed. (CIRCLE ONE NUMBER)

Grade School	High School	College/Technical	Graduate School
1 2 3 4 5 6 7 8	9 10 11 12	13 14 15 16	17 18 19 20+

This response format produced responses which could be coded as interval level data (with the possible exception of the last number of 20+, which was designated as an upper limit).

NEW QUESTIONS FOR APPROVED TOPICAL AREAS

It will often be necessary for the analyst to write one or more new questions for particular aspects of a given topical area found in the Compendium. For example, there are some flood insurance questions on pages 18 and 19 of the first approved questionnaire in the topical section of the Compendium titled Urban Residential Flood Damage. Questions in this section of the questionnaire ask if residents have flood insurance, the amount and kind, how much residents would pay for flood insurance, and the type and amount of flood losses covered. A recent Corps survey included modified questions on whether or not residents had flood insurance, what kind, and how much. In addition, two related questions were added to determine whether or not some residents had dropped flood insurance that they had purchased in the past, and the reasons for doing so. This was necessary to obtain the full range of flood insurance information desired. They are shown below:

Some people have had flood insurance policies that they have discontinued at some point in the past. Have you ever <u>discontinued</u> a flood insurance policy? (Circle Number)

1. YES   2. NO

If you answered yes above, why did you discontinue your policy? (Circle Number)

1. POLICY COST INCREASED
2. LOSS OF JOB/REDUCED INCOME
3. OTHER PERSONAL PROBLEMS
4. DISSATISFACTION WITH PAYMENT AFTER FLOOD
5. NO LONGER CONSIDERED FLOODING A SERIOUS RISK
6. OTHER REASONS (Please Specify): ________________

The first thing that must be taken into consideration when adding a relevant question like this is whether or not the desired question is within the "intent" of the approved questions and questionnaire components included in the Compendium. The above examples were questions needed for additional kinds of flood insurance information, a generic area of questions contained in the Compendium.

After deciding to add a question, the analyst should define it in terms of exactly what the question is to measure and what level of measurement is required. This must be done to produce valid and reliable information that is appropriate for analysis. It is accomplished by careful attention to question wording and formatting. Computer data entry considerations are also important when writing a question. Numerically coded responses are usually desired for analytical purposes, and restricting manual data entry to numerical codes can contribute to coding efficiency and the reduction of human error. After a question is written it must be pretested, and, depending upon results of the pretest(s), revised one or more times.

The two questions referred to above were designed as operational measures of the variable "flood insurance discontinuance" and the reasons for it. Both questions were designed to provide data at the nominal level of measurement. The numbered answer categories were used in analysis to produce nothing more than the percentage of respondents who said "yes" they had discontinued flood insurance; and, for those who said they did, the percentages for each of six reasons provided the possible answers to the second question.

The analyst worded both questions with the intent of being as straight forward and unambiguous (reliable) as possible, communicating to respondents exactly what information was being requested (validity). The answers provided were designed to be mutually exclusive (every answer completely different from every other answer) and exhaustive (all possible answers are provided). An "Other Reasons" category was listed as the last response to the second question to ensure that the responses provided were exhaustive of all those possible.

Both questions were written for a self-administered mail questionnaire and therefore the instruction "Circle Number" was written after each one in parentheses. The instruction "Please Specify" was similarly written in parentheses after the "Other Reasons" response to the second question. This was to inform respondents that the kinds of other reasons were to be written on the blank line provided. The line itself also helps to communicate to the respondent that an additional response is desired.

Both questions were formatted with answer categories numbered and indented so as to provide some white space around them to avoid a crowded appearance. Answer categories were typed in all capital letters to make them readily distinguishable from the questions themselves. The numbers in front of each answer identified the answer codes which were later typed into a computer data file for analysis.

Placement in the questionnaire is also an important formatting consideration. New questions should be placed in the questionnaire together with similar questions already there, with the more general and easier to answer questions placed first. Both of the new questions above were placed immediately after, rather than before, the other flood insurance questions in the questionnaire. The other questions were asking for the more general information about whether or not respondents now have flood insurance on buildings and contents, and the dollar amounts of such insurance. It is easier for the respondent to give these kinds of general answers about present insurance before answering the two more specific questions about past insurance.

CHAPTER IV

SURVEY SAMPLING AND IMPLEMENTATION

Survey design using the Compendium involves specifying the survey objectives, deciding how the survey is to be conducted (by mail, telephone, face-to-face, or some combination of methods), and then choosing appropriate Compendium questions. As described in the previous chapter, these questions must then be adapted to the survey objectives and the selected method of survey administration. Sufficient modifications are typically needed to make the questions fit together in a final survey form or instrument. An appropriate sample must then be designed and drawn prior to survey implementation.

Implementation transforms the survey instrument from a stack of blank forms to usable data, ready to code and analyze. Proper implementation is crucial to the success of any survey. The most wonderfully designed survey instrument can be unusable if implementation is faulty. Improper implementation can lead to an inadequate number of completed surveys, biased responses, an invalid sample, or unusable completed survey forms.

This chapter approaches survey sampling and implementation from the point of view of the survey nager. The problems of selecting a sampling frame, determining sample size, choosing a sampling method, and managing survey personnel are first addressed. The three major types of survey procedures are then described: face-to-face, mail, and telephone. The strengths and weaknesses of each method are discussed, guidelines are presented for each of these methodologies, and examples of each procedure are presented.

SAMPLING PROCEDURES[1]

Sampling is an extremely important and often neglected part of survey research. Samples should be randomly drawn from a carefully defined sampling frame. A sampling frame is a list of all those in a population who qualify to be interviewed for the survey. In a random sample from a particular sampling frame, every potential respondent in the sampling frame has an equal chance of being selected.

The size of the sample is often wrongly assumed to be the only criterion for judging the adequacy of the sample. It is often assumed that a large number of completed questionnaires are all that is required or that a large number is always better than a small number. In fact, the representativeness of the final, useable set of completed questionnaires is the most important characteristic of a sample. High non-response, failure to sample (represent) some sub-groups, sampling plans that give some sub-groups a higher probability of being surveyed (represented), and failure of some data collectors to follow sampling plans all threaten the representativeness of a sample. Random sampling

---

[1] Much of this discussion of sampling is derived from Chapter V of the Volume II Urban Flood Damage Manual (1991), "Designing and Drawing the Sample". The reader is referred to that manual for examples and for more detail than is presented here. Additional information is contained in a detailed discussion of sampling in Volume II of the Recreation Manual (1986).

from a carefully designed sampling frame is necessary to obtain a representative sample. This must be followed by maximum efforts to obtain survey data from every respondent included in the sample that has been drawn.

A self-selected sample, as opposed to a randomly selected or representative sample, is likely to be biased toward a particular view or condition. For example, if a post-flood damage survey allowed interviewees to choose whether they would receive a questionnaire or not, those who suffered large losses might be more likely to volunteer to be interviewed in hopes that the survey will somehow help them. Such a sample would very likely be biased toward those who experienced the most damage, and survey results, such as the average loss experienced or time needed for repairs, would be exaggerated.

DEFINING THE SAMPLE FRAME

The first step in the sampling process is to define the sampling frame. The sampling frame is the part of a population from which a sample is to be drawn. The sampling frame can be the entire population of a geographic area, but, for most Corps studies, the sampling frame is limited to only the households, individuals, businesses, or organizations that are relevant to the study purpose. A sampling frame for a flood damage survey, for example, is generally limited to those properties which the flood control project will directly affect.

DETERMINING SAMPLE SIZE

For certain studies, it may be possible to survey all population units (e.g. all adults or households) within the sample frame; but, for most studies, a statistically valid random sample is all that is practical or necessary to achieve a non-biased representation of the population. When all households, individuals, organizations, or other "units of analysis" in the sample frame are surveyed, the sample is called a census or 100% sample. It is usually only possible to collect survey data from all population units of analysis when the sampling frame is small. For example, a survey for the Montgomery Point Lock Study on the Arkansas River identified a relatively small sampling frame of shippers, terminals, and carriers of 120 entities. It was possible to survey all 120 entities of the entire sample frame with face-to-face interviews over a three-week period. In this case nothing would have been gained by taking a sample of less than 100% of the population, therefore a complete census was conducted. Costs, time, and logistical constraints, however, usually make it necessary to draw a sample of much less than 100% of the population.

When the sample frame is larger, the survey budget will usually not allow all potential units or entities to be sampled. It is then necessary to select a smaller number to represent all those in the sample frame. The basic formula for sample size required at various levels of precision is dependent on the variance of the most important variables to be used in the data analysis, and upon other factors affecting the complexity of the mathematical models to be used to fit the data. The sample size formula for the simplest univariate model used to fit the data is given below:

$$n = t^2 * \frac{s^2}{(\overline{Y} * r)^2}$$

where,

n = the sample size

s^2 = The variance (standard deviation squared) of the critical variable.

$\overline{Y}$ = An estimate of the mean of the critical variable, typically taken from a past study.

r = The level of precision desired (e.g., .05 or .1).

t = The t table value corresponding to the probability that the resulting sample estimate of the variable mean will be within the specified range of precision.

To use the above formula to calculate the required sample size "n" for a survey the analyst must first estimate the variance "s^2" expected from the survey for the variable of critical importance to the survey (e.g. flood damage repair costs to residential structures). Sample size calculations may be made for several different variables if more than one is of critical importance. The variance estimate used for "s^2" is a best guess estimate for the survey, usually the variance found in a similar study completed in the past. Another best guess estimate for the variance is the squared difference in the high and low values that might reasonably be expected for the variable, divided by four (Schaeffer, et al., 1979, p. 43).

The expected sample mean "$\overline{Y}$" of the critical variable for which sample size is being calculated must also be estimated. The best estimate is usually the sample mean found for the variable in some past study. The level of precision desired for this sample mean is designated by "r" in the above formula. The value for "r" that is used in the formula must be decided subjectively by the analyst. This is the number of plus or minus percentage points of <u>tolerable</u> error, away from the true mean of the variable for which the sample size is being calculated. The analyst simply decides how many percentage points of error in the resulting survey estimate of the sample mean are acceptable.

A "t" value must also be selected by the analyst using the above sample size formula. This value is selected from the t-table found in the back of most statistics books. The value of "t" selected should corresnond to the level of probability (from the table) that will provide the analyst with the desired level of confidence that the variable sample mean will be within the specified range of precision ("r"). For example, a "t' value of 1.96 would provide 95% confidence that the estimate of the mean is within the number of percentage points of the true mean specified by the level of precision. Other commonly used "t" values are 2.58 for 99% confidence and 1.65 for 90% confidence.

Note that for a small population, where the sample size is more than 5% of the sample frame, the same level of precision can be obtained with a relatively smaller sample size. This is estimated by multiplying the formula by a finite population correction (fpc) factor to derive the final sample size needed for such cases. With N being the population size and n being the initial sample size estimated by the above formula, the fpc is calculated by:

$$fpc = \frac{N-n}{N-1}$$

CHOOSING THE SAMPLING METHOD

A sampling method should be chosen which is most efficient for conducting a random selection of individuals, businesses, organizations, or other sampling units, and which minimizes the potential for sampling bias. Following is a brief synopsis of some of the most commonly used methods of sampling. For a more complete discussion of survey sampling, the reader is referred to the Volume II Recreation Manual (1986) and to standard survey sampling texts by authors such as Kish (1965) or Schaeffer et. al. (1979).

There are two ways in which random samples are most often selected, by simple random sampling or by systematic sampling with a random starting point. Other useful random sampling techniques include cluster sampling, multi-stage sampling, and stratified sampling. With each of these techniques, the actual selection of the sample units is usually by either the simple random or systematic method.

Simple Random Sampling. Simple random sampling is the most straight-forward sampling method. It makes no distinction for any sub-grouping of the population. It merely involves making random selection of potential respondents. The random selection can be accomplished by assigning a number to each potential respondent or sampling unit and using a computerized random number generator or table to determine which ones are to be included in the sample as survey respondents.

It is usually first necessary to list all units in the sampling frame (e.g. names, houses, or businesses), making sure that every appropriate unit is on the list but that none are included more than once. All sampling units included on the list are then numbered, starting by assigning a number one to the first unit in the list. The size sample desired is then drawn by selecting from the pool of numbers assigned to the list, using a random number table or a computerized random number generator.

For very small populations, using manual techniques for randomization may be just as efficient as using a random number table or computer. For example, all potential respondents from a population of 100 or less could easily be named on individual three by five cards. The order of selection from this deck of cards could then be randomized by simply shuffling the deck. The size sample desired could then be selected by simply drawing the required number of cards from the top of the shuffled deck.

Systematic Sampling. Systematic sampling involves putting the potential respondents in random lists and choosing a random number to select the first respondent. Thereafter, every nth member of the sample frame is selected, where "n" is that portion or fraction of the sample frame needed to draw the required sample size from the entire list. For example, suppose the analyst decides to systematically draw a sample of 100 urban residences from a population of 400 residences which have been flooded. All 400 residences would first be listed in numerical order from 1 to 400, and "every nth" for a systematic sample would be every 4th residence on the list. The order of the numerical listing could be arranged alphabetically by property owners' last names, by the street number, or by some other method which could be assumed to produce an unbiased listing. The list should not be ranked in any way. Ranked or periodically occurring units on the list with similar characteristics can result in sample bias.

To choose a starting point on the list for drawing the sample, a number between one and four would be randomly selected (e.g. by shuffling four numbered cards and drawing one from the top of the shuffled deck). A number between one and four is used in this example because the desired sample size of 100 is one fourth the total population of 400, therefore the sampling interval is "every 4th". If, for example, number three was on the card randomly chosen from the shuffled deck, then residence number three in the list of four hundred would be the starting point for the sample - the first residence to be selected. The next residence selected would be number seven, then number eleven, and so forth until "every 4th" residence in the entire list is selected for a total sample of 100.

Unbiased systematic sampling assumes a randomly constituted listing of the population free of any kind of periodically occurring characteristic. If no such list exists and it is not possible to develop one, a systematic sample may be biased. For example, periodic occurrence of houses located on block corners in a listing of city residence addresses could bias a sample. If the interval between these corner locations coincided with the sampling interval used with such a listing, and if the starting point for the sample also happened to be a corner location, then only corner house locations would be included in the sample.

Cluster Sampling. Cluster sampling is done by taking a random sample from a sampling frame by randomly selecting clusters of potential respondents as sampling units rather than selecting each respondent individually. Clusters usually are geographically based, such as neighborhoods or city blocks. Clusters can also be based in time. For example, random samples can be taken of individuals visiting a recreation site by randomly selecting clusters of time, and treating all individuals encountered during the selected time periods as respondents.

Cluster sampling can be a time and money saver, but clustering does introduce additional sampling error. Its benefits have to be weighed against this potentially important disadvantage.

Multistage Sampling. Multi-stage sampling is similar to cluster sampling because at the "first stage" large groupings or clusters of potential respondents are selected. At the "second stage", a random sample is made of each cluster selected at the first stage, and so forth for subsequent stages. Multi-stage sampling is more applicable to large geographic sample frames where face-to-face surveys are to be used. It can be a cost saver when time and available funding do not permit interviewing potential respondents in all parts of a large geographic region.

Stratified Sampling. If any of the analysis is to be done on classes or sub-groups of data, it is important that all of the relevant strata (sub-groups) are represented. If the sample is relatively small, or if certain groups are expected to be under-represented in the sample frame, a stratified sample is often recommended. Stratifying a sample also serves to eliminate or minimize sampling error for the categories or sub-groups of particular variables. For example, in stratifying a sample on the variable "place of residence", when it is known that 50% or the residences are urban and 50% are rural, exactly 50% of the sample is drawn from both urban and rural areas respectively. A stratified sample assures that each stratum or sub-group (e.g. urban vs. rural) is sampled in proportion to the total population of that stratum. Stratified sampling breaks the sample frame into these strata and random sampling is conducted proportionately for each stratum.

It is highly recommended that survey research and statistics text books be consulted for more details on sampling. Among the texts which could be consulted are How Many Subjects: Statistical Analysis in Research, by Kraemer and Thielmann (Sage, 1987), Survey Sampling, by Kish (John Wiley, 1965), and Elementary Survey Sampling, by Schaeffer, et. al. (Duxbury Press, 1979).

Further description of each of these sampling strategies can also be found in the Volume II Recreation Manual (1986), and in the Volume II Urban Flood Damage Manual (1991).

Implementing the Survey Sample. Implementation of sampling is done by strictly adhering to the list of population units selected for the sample. Not everyone in the sample of respondents selected will be available or willing to participate in the survey, but it is still necessary to follow only the list of selected sample units. Units not included in the sample should not be surveyed.

If a survey is well-publicized, or if news of it spreads by word-of-mouth, there are likely to be a number of people who will request to be surveyed. As inviting as it seems to get eager and willing interviewees, these people should be told interviews can only be done with people selected in the sample, otherwise the sample may be biased. In cases where people outside of the sample insist on being interviewed, either because they have suffered a very severe flood loss or for some other reason, it is permissible to interview them if their completed questionnaires are kept separate from other data during analysis. Such unsolicited questionnaires should be reported the same way that unsolicited letters are reported; receipt of them is acknowledged but their content is not usually analyzed.

PRETESTING

The survey questionnaire, the sample listing, and the method of survey administration should all be thoroughly pretested before a survey is implemented. This usually involves several repeated cycles of pretesting, corrections, additional pretesting, and more corrections until the analyst feels the survey is finally ready to be implemented.

PRETESTING THE QUESTIONNAIRE

The questionnaire should be pretested with a small group of respondents who can be debriefed at the end of the pretest interview. Questionnaire pretesting is conducted to detect all possible problems. The objective is to identify things such as poor wording or sequencing of questions, inadequate question response options, and questions which respondents refuse to answer. Problems can be detected by observing respondents as they complete a questionnaire, reviewing their responses (or lack thereof), and asking them to comment on any problems they noticed while responding to the questions.

PRETESTING THE SAMPLE LISTING

Pretesting the sample listing to be used for the survey is important to assure that selected respondents included on the list can be contacted. It is assumed that every name selected in the sample can be contacted and asked to participate, but this is very often not the case. Severe sampling bias can result from large numbers of bad addresses for mail surveys, or of bad phone numbers for phone surveys.

The use of phone directories to identify the names and addresses of respondents for mail surveys can cause problems due to insufficient address information. Residents living in rural areas may have only the name of their nearest rural community listed. Using such incomplete addresses can result in large numbers of mailed questionnaires which are undeliverable. One alternative to telephone book listings for mail survey samples is use of automobile registration listings.

In addition to sometimes containing incomplete addresses, telephone directories also can quickly become outdated. When phone directories are used for sampling, the analyst should be sure the most recently published version is used.

Another potential problem with some telephone directories is omission of a large proportion of the population due to unlisted numbers. For large cities and certain parts of the country, a relatively large number of residents with telephones (often 20% or more) are not listed in the telephone directory. Use of such a directory as a sampling frame can result in a biased sample.

An alternative to using directories for telephone surveys is random digit dialing. This requires identifying the existing blocks of telephone numbers in service for a region and randomly sampling all of those numbers. Random digit dialing gives everyone with a telephone a chance to be included in the survey. Random digit dialing may require many extra calls, because many phone numbers in service for a region may not be included in the analyst's sampling frame (e.g. business phone numbers contacted during a survey of residential property owners).

Commercial survey sampling companies maintain nationwide computer files of up-to-date telephone listings, from which they draw random samples from specific geographic areas for sale to researchers. It is also possible to purchase random digit samples of phone numbers from them for specific regions of the country.

PRETESTING THE METHOD OF SURVEY ADMINISTRATION

The method of survey administration selected should also be pretested to make sure that it works effectively. This is particularly important for mail surveys. Bulk mailings conducted to reduce postage costs often result in more undeliverable questionnaires, particularly if addresses from the sampling list are not as complete as they should be. Bulk mailings are also often irregularly timed because they are treated as low priority by post offices. Respondents may also be more apt to respond if first class stamps are used on mail-out and mail-back questionnaires. These and similar details should be pretested by a small sample mailing, followed by a phone debriefing of those to whom the sample mailing is addressed. They should be asked if and when they received the mailed survey materials, and about their general reactions to the study purpose. The questionnaire and cover letter should also be assessed.

PERSONNEL MANAGEMENT

No matter what survey method is used, it is necessary to have someone oversee the entire survey process who is completely familiar with the survey objectives, the survey instrument, and all of the details of the implementation process. It is best if this principal investigator is involved in the actual design of the survey. Previous survey experience is important for knowing how to address the many pitfalls which might disrupt the survey process, such as unhappy interviewers, angry interviewees, or a sampling procedure which has gone awry. While most Corps surveys can adequately be supervised by one person, it is important to keep in mind that an adequate ratio of supervisors to interviewers or mail survey workers is very important. No supervisor should be expected to manage more than eight to ten workers at a time.

One of the most important jobs of the survey supervisor is to ensure that all the survey forms are adequately completed. This is best done when the interviewers are easily accessible and the interviews are fresh in the interviewers' minds.

FACE-TO-FACE SURVEYS

Face-to-face surveys offer an opportunity for the highest proportion of completed surveys. People are less likely to refuse interviewers who have taken the trouble to come to their home, office, or other survey location (e.g. recreation area). The face-to-face interview also allows for more in-depth, complex questioning, as well as questions that require flash cards or other visual props. Respondents may be less likely to skip individual questions if an interviewer is on the scene to ask the questions and probe or encourage response.

The personalized aspect of the face-to-face survey, however, also lends itself to the possibility of interviewer bias and other types of interviewer error. Interviewers can unintentionally influence response through unguarded non-verbal messages. The first time any particular question is asked, it should be asked exactly as it appears on the interview form. The interviewer should realize the wording of each question was carefully chosen to obtain very specific information. Changing even a few words can significantly change the meaning of the question. There are two exceptions where the wording of a question may be altered. One instance is where an individual respondent has difficulty understanding the meaning of a question and the interviewer is able to rephrase the question in a manner that does not change its meaning. The interviewer must be very familiar with the survey instrument and aware of the meaning intended by the question. A second exception would be when problems have occurred with the wording of a question and the interview supervisor changes the wording. In this instance, the wording should be changed for all interviewers, so that the respondents are answering the same question.

INTERVIEWER SELECTION

For face-to-face surveys, field interviewers must be hired, trained, scheduled, and their completed interviews checked. Interviewer transportation, food, lodging, and safety must also be coordinated.

Interviewing requires no specific academic background. Any intelligent, friendly person with a positive attitude can be trained to interview. However, adequate training is essential even for interviewers with experience from previous survey projects. The primary qualifications of an interviewer are not to be shy about meeting people and asking, sometimes personal, questions; to be willing to present a well-groomed, non-threatening appearance and a non-threatening demeanor; and to be able to ask questions as they appear in the questionnaire without injecting any personal bias or excess conversation. An interviewer must be "thick-skinned" enough so that refusals or hostile respondents do not cause him or her to react with anger. Anger is not permissible, no matter what situation may arise. On most days he or she will likely experience the utmost cooperation and people more than willing to answer every question. However, the interviewer must be equally prepared for those days where people don't answer the door, skip appointments, refuse to answer questions or are simply hostile.

INTERVIEWER TRAINING

Training sessions are an absolute necessity. Even the most highly skilled and experienced interviewers need some orientation to the questionnaire and specific survey procedures that are to be employed. From one to two days should be allotted to training, depending on the complexity of the survey instrument and the experience of the interviewers. The training should include an explanation of the intent of the overall survey, a description of the sampling process, a discussion of how to contact interviewees, the importance of not biasing the interviewee when asking the questions, how to respond to various situations during an interview, and making sure that a questionnaire is completed in a legible manner so that it can be properly coded.

MONITORING INTERVIEWERS

The job of the survey manager continues to be very important after the interviewers are trained and begin to work. The survey manager usually assumes the role of the primary public contact during interviews. The manager makes periodic checks of completed survey questionnaires to see that questions are answered adequately and that the handwriting of interviewers is legible. In a face-to-face survey it is valuable for the survey manager to attend at least one survey session with each interviewer in order to determine that the interviewer is asking the questions as they appear in the questionnaire. These monitoring activities are very important for minimizing human error and maximizing the quality of survey results.

TELEPHONE SURVEYS

Telephone surveys are the least time-consuming survey to undertake. They require no travel, generally no appointments, and there is no waiting for surveys to be returned in the mail. Even if the interviews are all long distance calls, the savings in time and travel may easily more than offset any telephone charges.

Telephone surveys can be facilitated by the purchase of phone numbers from a sampling firm. The specific area can be identified, down to zip code level, and a random sample of phone numbers, complete with accompanying names and addresses, can be acquired. The addresses can then be sorted to identify the properties in a smaller geographic area.

One difficulty with telephone surveys is the large percentage of unlisted phone numbers throughout the United States, approximately 28 percent nationwide and over 50 percent in some metropolitan areas (Survey Sampling, Inc. 1988). Without these unlisted numbers, a significant portion of the sampling frame, generally wealthier people and people in certain professions with a need for unlisted numbers, would be under-represented. Telephone numbers can be obtained for unlisted numbers, but the only geographic specificity is the area covered by the telephone exchange. As previously explained in the discussion on pretesting, random-digit dialing can be used to obtain random samples of telephone numbers for populations of residents for whom a large proportion of the telephone numbers are unlisted.

Response to well designed telephone interviews is generally very good. Response rates may be lower in certain areas of the country, which may, for example, be subject to heavy tele-marketing.

INTERVIEWER SELECTION, TRAINING, AND MONITORING

As with face-to-face interviewers, telephone interviewers must also be hired, trained, scheduled and have their completed interviews checked. In addition, the administrator in charge of a group of telephone interviewers should periodically monitor on-going telephone interviews to make sure they are being correctly conducted. Complaints from respondents must be resolved and special call-backs scheduled for those who do not speak English or require special attention for some other reason.

CATI SYSTEMS

Telephone interviewers often use a system called Computer Assisted Telephone Interviewing (CATI). The CATI system allows for the telephone interviewer, equipped with a computer, to work directly from a script that appears on the computer monitor. Answers can be entered as they are given by the respondents. If there are logical skips in the sequence of questions, based on the answers given to certain questions, the computer will automatically proceed in the logical order. For instance, suppose there are a series of questions regarding a flood warning message and the respondent answers "no" to the first of these which asks whether or not the warning was heard. When the interviewer enters this "no" answer, the CATI software will skip over all the other questions asking about the flood warning and go directly to the next series of questions in the survey.

MAIL SURVEYS

The mail survey is the least expensive means of reaching large numbers of people. Ideally, it would merely involve taking a random sample of an address list, doing a single mailing with a cover letter and waiting for everyone to respond by return mail. A mail survey, however, is a lengthy process involving a series of mailings, as described below. A mail survey offers the advantage of providing the respondent with maps and illustrations than can allow for more informed responses. If not well done, however, mail surveys can be subject to very low response rates (sometimes less than 10 percent) and unrepresentative response.

Several rules of survey design have made it possible for mail surveys to have a substantially greater response. These rules follow the "total design" method, devised by Don Dillman (1978):

1. Participation in the survey can be enhanced by pre-survey publicity, such as press conferences and news releases. This will make people aware of the importance of the survey and its legitimacy.

2. Pre-survey phone calls can serve as a personal solicitation to participation in the survey. These calls may be an added early expense, but may save in the long run by reducing printing, mailing, and labor costs. Phone calls will also increase response by obtaining commitments from some people who would not ordinarily agree to complete the survey. It is important that the pre-survey screening be honest about the nature of the questionnaire and the amount of time required to complete it. Otherwise, people may feel annoyed and be less likely to respond.

3. Include a persuasive cover letter that emphasizes the importance of the respondent's participation in the survey. The letter must state the voluntary nature of the survey and assure the respondents as to the confidentiality of their responses. Incentives can be offered to help increase the response. For

instance, copies of Corps recreation maps or a directory of local community services could be included in the mailing. A mail survey in Luzerne County, Pennsylvania, and Orange County, California offered respondents their own copy of a household inventory survey. By tearing off perforated carbon copy questionnaire sheets, respondents could retain a copy of their survey inventory for their insurance records.

4. The survey form itself is most effective when it takes the form of a small (5 inch x 7 inch), attractive booklet. The cover of the booklet should be made of heavy card stock and should be illustrative of the project purpose. The survey form should have very simple, straight-forward instructions. Instructions can also refer the respondent to other information, such as insurance records, which may aid in completion of the questionnaire. If necessary, maps and project illustrations can also accompany the mail survey questionnaire.

5. Always make it easy to respond. The survey should always be accompanied by a self-addressed, stamped envelope or be configured as a self-mailer. The respondent should not be expected to provide the postage or envelope. If there is concern that respondents might be too worried about their privacy to participate, respondents can be promised that no identifying numbers will be put on the questionnaire and asked to send in a separate post card with their name on it to show that they have returned their survey form. Receipt of the post cards then

becomes the only way for the analyst to know which persons from the sampling list have responded to the survey.

6. Follow-up the existing survey with a reminder post card about a week after the initial mailing. If there is still no response, a second survey form should be mailed out two weeks after the post card to those who have not responded.

Timing of the steps involved is very important and it is important to see that personnel adhere to the schedule. When questionnaires are returned, the names and addresses of respondents must be immediately eliminated from the survey sample list so that they are not included in subsequent mailings.

In a mail survey, the survey manager's phone number should be in the initial cover letter or on the survey form, along with the name of a contact from any local sponsoring organization. This makes it possible for respondents to phone if they have questions about how to respond or doubts about the legitimacy of the survey. The mail survey manager also fills the role of a scheduler of mailings, including the initial mailing, follow-up post cards, and follow-up mailings of the questionnaire.

EMPLOYEE SELECTION AND TRAINING

Individuals involved in mail surveys have much different responsibilities from those conducting either face-to-face or telephone interviews. Except for a possible telephone contact to check on non-response or for some other reason, mail survey workers have much less contact with the public. These workers do not need to be outgoing or resistant to public criticism.

Instead, they need to have patience and capacity for detail in an involved process of organizing mailing material, stuffing envelopes with appropriate material, and recording and filing the questionnaires as they are returned.

PROCEDURAL GUIDE TO IMPLEMENTATION

Every survey can benefit from a comprehensive procedural guide. The guide should describe the survey objectives, the objectives for each section of the questionnaire, and provide detailed procedures for conducting the survey. This guide gives a continuity to the survey process which is particularly important if a key individual is no longer involved. It can also serve to document the survey process.

One section of the Compendium includes instructions for the actual field administration of the example survey questions. Those instructions are in Section I, for Dock and Carrier Survey instruments. Many of the general components of the Dock and Carrier Survey instructions may, in a generic sense, also be appropriate for other surveys. For example, reference should always be made in the instructions to <u>who</u> is to be interviewed (the population or sample) and <u>how</u> initial contact is to be made with the selected respondents. However, the example in Section I will, in most cases, need to be revised in many ways for each new dock and carrier survey. As such, it should be treated as a generic example upon which to build.

GENERAL INSTRUCTIONS

General instructions should be prepared to orient the interviewer to who, within the constraints of the sample, qualifies to be interviewed and who does not. General instructions should emphasize the importance of adhering to the sampling list and not making any substitutions. Emphasis should also be placed upon the importance of legibly recording all data and being sure to ask and obtain answers to all questions, except for cases when respondents object.

General instructions should repeat training principles for how interviewers should introduce themselves and the study to respondents. The interviewer should be reminded of the importance of establishing rapport with respondents and getting a sincere expression of willingness to participate in the survey before beginning to ask questions. A confidentiality statement is also normally included with these interviewer instructions. Respondents should be assured that responses to survey questions will be reported only in aggregated form. They should also be assured that their names, addresses or other personal identifiers will not be associated with their answers to survey questions. Their names, addresses, and other personal identifiers should be separated from the response file once a data base is created.

General instructions should also alert interviewers to the importance of recording all miscellaneous comments from respondents. These often disclose reasons for non-response early in a survey which can be corrected as the survey progresses.

SPECIFIC INSTRUCTIONS

Specific instructions should be prepared for particular questions for which it can be determined (perhaps through a pretest) that some respondents may need clarification, assurance of confidentiality, or an explanation of why the particular information asked for is needed. It is sometimes possible to prepare a list of standard responses for interviewers to give to respondents in reply to requests for clarification and other information.

Face-to-face and telephone interviewers need specific instructions to "probe" and provide "feedback" on certain questions. Probes are particular words, phrases, or sentences which interviewers can use to elicit more complete responses when respondents do not initially provide all information requested. Feedback phrases are used as positive cues or verbal rewards to respondents at critical points in the interview, indicating to them that their efforts in responding are appreciated. This can be important for successful completion of long interviews, but should not be overdone for fear of biasing response.

Interviewers may also need specific instructions for administering willingness-to-pay or contingent valuation questions, so as to guard against "strategic bias" on the part of respondents. Answers reflecting strategic bias are values which are purposely inflated or deflated in hopes of influencing the planners, managers or policy makers who may use the survey results from these questions.

CODING INSTRUCTIONS

It is critical to have a detailed, written set of coding instructions, sometimes called a "code book", so that there is no ambiguity about the way in which responses to all questions are to be coded. Codes for missing data and instructions for imputing responses for some kinds of missing data must be made explicit. A specific code such as -99 should be used to indicate missing data values. Blanks and periods are also read as missing values by some computer statistical packages, but this can cause problems if at some point in the analysis the data must be transferred to another statistical package or data base which cannot accommodate blanks or periods.

CHAPTER V

DATA EDITING AND ENTRY

Screening survey forms, data entry, and data cleaning are all necessary before any data analysis can begin. This section describes what is undoubtedly the most tedious part of the survey process, especially when there are many long survey forms to code. However, numerous, irreparable errors will occur when these very critical tasks are not carefully performed. Therefore, the resources devoted to these task should be substantial.

CLERICAL EDITING

DETERMINING IF A SURVEY FORM IS USEABLE (ROUND ONE)

Before taking the effort to type the information from a completed survey form into a computer data base, it should first be determined whether the survey form is useable. Screening at this point is usually not a matter of how reasonable the answers are, but a matter of completeness, legibility, and whether basic recording instructions have been followed. At the end of this chapter the issue of the usability is revisited, based on computer checks for the completeness and consistency of the answers.

<u>Completeness.</u> Completeness can be determined by a specific established criterion. For instance, the analyst can establish that a certain percentage of answers must be filled in before a survey form is considered complete enough to be usable. There

may also be a number of critical questions that must be answered before a survey form is considered useable. For example, if the only objective of a particular survey is to determine depth-damage functions and no damage information is entered, then that survey would be unusable, even if all the background and value questions were completed.

Legibility. An important criterion of useability is whether or not a survey form is legible. One must be careful not to introduce bias when interpreting ambiguous markings. If there is any doubt and the person who completed the survey is unavailable, then the answer should be considered missing.

Following Instructions. Another important consideration is whether or not the coding instructions have been followed. A typed "coding book" should be prepared which gives directions to interviewers and data entry personnel on the location and format of each variable. Self-administered questionnaires are particularly subject to data entry errors by the respondent not following instructions and from data entry personnel incorrectly interpreting responses. These questionnaires always have to be reviewed in detail. If answers are not marked in the correct location or in the proper format called for in the coding book, such as the correct alpha or alpha-numeric code, then consideration should be given to deleting or attempting to correctly code that answer. If much interpretation or guessing is required to determine what answers are intended throughout a questionnaire, then that survey form should not be used.

CORRECTING RECORDING PROBLEMS

Any of the problems described above can best be prevented in face-to-face and telephones surveys by having a supervisor inspect every survey form as it is submitted and correct the problem immediately. It is ideal if interviewers are available during data entry to help interpret any ambiguous survey responses. Immediate attention to response problems is particularly important when interviewers are working under contract and will no longer be available when the interviewing is completed. If these problems remain, and the interviewer is not available or is unable to recollect what the intended answer was, it is better to consider an answer as missing than to rely on guesses of what the respondent intended. Guessing is likely to introduce error in the data.

DATA ENTRY

The efficiency of the coding operation is very dependent on the quality of questionnaire design. An inefficient design can confuse the person coding the data, slow the coding process, increase study costs, and lead to coding errors. Not only is it important that the form be well designed but also that the forms be neatly filled out according to survey instructions.

INTERVIEWER CODING AND DATA ENTRY

Computerized direct data entry systems allow for interviewers to input the answers they receive into a computer file while an interview is being conducted. These systems have existed for some time for telephone interviews, and direct data entry has more recently been developed for face-to-face surveys.

A direct face-to-face interviewer coding system was initiated in 1990 by the Corps of Engineer's Waterways Experiment Station (WES). One of the initial applications was on a traffic stop survey of lake recreation users for the Visitation Estimation Reporting System (VERS). Interviewers at Corps lakes were equipped with portable computers. The computers contained direct data entry system (DDES) software. The DDES allowed for questions to appear on the computer monitors just as they would look on a survey form. In the VERS application, interviewers stopped cars as they left recreation sites and asked for participation in a two-minute survey. There was little refusal and the survey format enabled interviewers to quickly enter the data so that no extra time was required of the respondent.

Some computer statistical packages are now equipped with data entry systems that would allow for the same direct data entry as was done on the VERS study. These systems are often equipped with data screening capabilities that allow limits to be placed on logically "acceptable" answers and allow for logical skips when a particular series of questions is only applicable to certain respondents.

SCANNING

Scanning is the process where hard copies of information can be read directly into a computer file by an optical device. The use of scanners in survey research is limited to closed-ended questions, where a response does not have to be written, but instead is selected from a list that has been provided. Since scanning is used to identify darkened spaces on scan forms, it has limited application when handwritten open-ended responses are required.

SPOT CHECKS OF CODING

Before final editing, it is advisable to spot check the data entry for all survey questions. This will allow the identification of any systematic problem with either the questions or the data entry. This procedure will assist in focusing the data cleanup effort.

COMPUTER EDITING

Data entry is only one step in establishing a computer file of the survey data. A data base format must be established either in a data base management system or a statistical package. The names, lengths, and position of each variable are identified and whether a particular variable is in a numeric or alpha-numeric code. Once the data base has been read into this new format, new variables can be created by mathematical manipulation of the initial variables.

SETTING LIMITS FOR OUTLIERS

All questionnaire items, except for totally open-ended questions, can have limitations placed on acceptable answers. Closed-ended questions, including dichotomous (e.g. yes-no) and multiple answer response possibilities, are limited to the answer alternatives supplied on the survey form.

While many computer statistical packages have methods for identifying outliers or extreme values for variables, there is no standard statistical rule or test on when outliers should be eliminated. It is generally left for the analyst to determine what is reasonable for any particular variable.

To facilitate identification of cases that might be outliers, the analyst can print all the cases that fall outside of a predetermined range or beyond a predetermined number of standard deviations from the mean. Prior to establishing a criterion for defining outliers, the analyst may want to examine the frequency distribution for each variable and a plot of the data.

CHECKING COMPLIANCE WITH FILTER QUESTIONS

Filter questions are designed to allow for built-in logical skips of survey questions that do not apply to a particular respondent. The following example of a filter question was used in a series of questions concerning flood warning from a survey in Houston, Texas:

> *Just before the <u>first flood</u> that affected you in 1989, did anyone at this residence hear from anyone or receive any other communication that flooding was possible? (circle number)*
>
> *0. NO if No, SKIP to Q24*
> *1. YES*

This question allowed anyone that missed the flood warning to skip to the next series of questions.

As mentioned above, filter questions are used to determine if it is appropriate to go on to the next group of questions.

They are an efficient way to check for inconsistencies. If no screening procedures are built into the data entry system to ensure the question filter process is followed, it is then necessary to build in other logical checks.

DETERMINING IF A SURVEY FORM IS USEABLE (ROUND TWO)

Computer screening allows a second, more efficient, opportunity to screen the individual forms after all data have been entered. After data entry, variable creation, re-coding, and initial computer screening, a decision should be made as to whether each particular case meets whatever requirements the analyst establishes for a survey form to be considered complete. Again, the analyst may want to establish that there are certain pivotal questions that must be complete or a certain percentage of the questions overall that should be complete before the survey form is considered useable. Computer checks can also be made to determine if any particular survey forms have an inordinately high number of extreme values. Those cases may also be candidates for elimination.

CHAPTER VI

DATA ANALYSIS

There are numerous types of statistical analyses that may be performed on survey data. Every statistical technique has specific assumptions which must be satisfied for the technique to be valid and the results of the analyses to be considered reliable. The principal objective of this chapter is to orient the reader to the nature and scope of data analysis, with particular emphasis upon some commonly used statistical tests and analysis procedures. This manual does not attempt to present a comprehensive discussion of all statistical techniques available, but rather presents guidelines to using some of the more common techniques. This discussion is supplemented by Appendix A, containing definitions of some common types of these analyses.

The chapter begins with the steps of statistical hypothesis testing. Levels of data measurement are then defined. This is followed by a description of sample selection bias and the possible need to weight data to correct for this bias prior to performing any statistical tests. Next is a discussion of some alternative ways to treat missing data. The following types of statistical analyses are then described: univariate statistical procedures (one variable at a time), commonly used bivariate (two variable) relational analysis procedures, and multivariate (three or more variable) procedures with special emphasis on regression analysis.

PERFORMING STATISTICAL ANALYSIS

It is extremely important to carefully plan the strategy for the analysis before the first piece of data has been collected. This is necessary to assure that the amount and types of data collected will be sufficient for the desired analyses.

DESCRIPTIVE PROFILES

In many surveys, the only data analysis conducted is the generation of descriptive profiles of responses to individual survey questions. This usually does not require any type of statistical testing.

Even if further analysis is to be conducted (e.g. hypothesis testing), it is useful to generate descriptive profiles as an important first step. These profiles of question response can help the analyst identify extreme responses and bad data, which can then be corrected or removed prior to further analysis.

HYPOTHESIS TESTING

Most survey data analysis also involves some hypothesis testing, even if not explicitly stated. For example, when estimating potential residential flood damage by comparing home contents values from a survey, an implicit hypothesis might be that the average value of contents of homes within the 100 year flood plain is lower than that of homes in the 500 year flood plain. Research hypotheses may be stated in various ways to accommodate a particular research situation. This chapter provides several examples of research hypotheses, but analysts should tailor their hypotheses to the needs of the specific analysis at hand. Hypotheses are tested by performing statistical tests. The steps in statistical hypothesis testing are as follows:

1. Formulate the question to be investigated in statistical terms. This is called the **alternative hypothesis, H_a**. For example, we may theorize H_a: The mean August 1991 cost of shipping on the Columbia River (μ_A) is <u>higher</u> than that on the Tennessee-Tombigbee (μ_B), or in statistical terms: $H_a:\mu_A>\mu_B$ where μ_A = mean August 1991 Columbia River shipping costs and μ_B = mean August 1991 Tenn-Tom River shipping costs. We could also have theorized that the mean cost of shipping on the Columbia River is <u>lower</u> than that on the Tenn-Tom, which, in statistical terms would be $H_a:\mu_A<\mu_B$.

2. Formulate the **null hypothesis**, H_O. To verify our alternative hypothesis (H_a), we will attempt to disprove the null hypothesis (H_O). For example, the null hypothesis (H_O) to an H_a shipping cost hypothesis would be that mean shipping costs are equal on the Tenn-Tom and Columbia River Systems (H_O: $\mu_A = \mu_B$).

3. Set the level of significance or **alpha** level (α) and the sample size (N). Alpha equals the probability of committing a **Type I Error**. A Type I Error results when the analyst rejects the null hypothesis when it should be accepted. A Type I Error is committed if we conclude that shipping costs on the Columbia River are higher, as hypothesized, when they really are not higher. This is in contrast to the **beta** level, which is the probability of committing a Type II Error, resulting when the null hypothesis is accepted but should be rejected. The alpha level and beta level are related in that as one increases the other decreases, though this is not a strict linear relationship. A Type I Error is committed if we conclude that shipping costs on the Columbia River are higher, as hypothesized, when they really are not higher. The seriousness of making a Type I Error should determine the choice of the alpha level. Although it is commonly set at α = .05, this doesn't have to be the case. Setting the alpha level at .05 means that we would tolerate no more than a 5% chance (resulting from a statistical test) of rejecting the null hypothesis incorrectly when it should be accepted (Type I Error). For critical decisions a more stringent level of alpha may need to be set. For example, if a decision to initiate a one hundred million dollar project on the Columbia River is dependent upon whether or not it has lower shipping costs, a lower level of alpha (.01 or less) would probably be more appropriate since the consequences of making an error in confirming reduced shipping costs would be quite serious.

4. Select the appropriate statistical analysis. Discussions of common alternative methods of univariate, bivariate and, multivariate statistical analysis are provided on pages 73-108 of this chapter.

5. Design the survey, collect a random sample, and prepare the data for analysis. These tasks are described in Chapters II - V.

6. Perform the statistical analysis. Details concerning this step will be provided in the remainder of this chapter.

7. Either reject H_o or fail to reject H_o based upon the results of the statistical analysis. The final result sought in most statistical tests is the **p-value**. This statistic is produced in the standard output for most statistical packages. Prior to the availability of computerized statistical packages, the researcher referred to standard statistical tables compiled for a specific test statistic to find the p-value corresponding to the resulting value of the test statistic. The p-value is the probability of observing a sample outcome as extreme as the observed result if H_o were true. A small p-value indicates that it is very unlikely we would have gotten the sample result we did if H_o were true, therefore, the null hypothesis must not be true. In practice, if the resulting p-value is equal to or less than the level of alpha set by the analyst in step #3 above, the null hypothesis (H_o) is rejected and, by implication, the research hypothesis (H_a) is confirmed. The p-value represents the weight of the evidence for rejecting H_o. The lower the alpha level set in step #3, the greater the evidence required (lower resulting p-values) to reject the null hypothesis. Setting the alpha level very low (e.g. .001) and finding a p-value resulting from the analysis that is equal to or lower than alpha indicates it is very unlikely the sample result would have occurred if H_o were true, therefore H_o should be rejected. Most researchers normally require a p-value of <.05 (the level of alpha commonly set as the minimum criterion of statistical significance) as sufficient to reject H_o. However, as explained in step #3 above for situations where the consequences of making a Type I Error are more serious, a lower alpha level may be set, which requires a correspondingly lower p-value resulting from the analysis to conclude statistical significance (rejection of H_o). A conclusion of statistical significance implies that the sample results can be generalized to the population from which the sample was drawn.

LEVELS OF MEASUREMENT

Before proceeding with a discussion of types of statistical analyses, a few basic terms must be defined. Data consist of measurements of one or more **variables** which are characteristics or attributes of the subjects in the study. A **case** refers to a set of characteristics or variables for a subject or respondent measured in a questionnaire. Each question asked by the researcher may be considered a variable.

One of the most important determinants of the appropriate statistical technique to select is the **level of data measurement.** Statistical tests may be grouped according to the level of measurement of data and the type of research question being investigated. There are four levels of measurement of data: **nominal**, **ordinal**, **interval**, and **ratio**, in ascending order. The level of data measurement for a variable will determine which statistical analyses are applicable in a given study situation using that variable. Variables measured at the ratio level are considered to be of the highest level, whereas nominal level data are at the lowest level of measurement.

NOMINAL-LEVEL MEASUREMENT

Data are measured at the nominal level when each value is a distinct category, i.e. a specific value serves merely as a label or name. No assumptions of ordering or mathematical difference between categories is implied. Variables are non-quantitative (e.g., names of streams or rivers, sex of respondent).

ORDINAL-LEVEL MEASUREMENT

With ordinal level data it is possible to rank-order the categories, but no mathematical property of distance between categories exists. A Compendium question rating effectiveness of a flood preparedness plan (Topical Section A, page 11) provides a good example of ordinal level measurement. There are four possible rating responses to this question: Excellent, Good, Fair, Not Effective. For analysis purposes they would usually be numbered 4, 3, 2, 1 respectively, with higher numbers indicating higher levels of effectiveness. However, an answer of "Excellent" (#4) cannot be interpreted to mean being twice as effective as an answer of "Fair" (#2). For some respondents, "Excellent" may mean ten or more times as effective as "Fair". The analyst can only be sure of the meaning of the order of the responses, that higher numbered responses indicate higher levels of effectiveness. Ordinal responses such as this do not indicate how much higher in effectiveness one possible response is compared to another.

INTERVAL-LEVEL MEASUREMENT

In addition to having rank ordering of categories, interval level data have the distances between categories defined in terms of fixed units (intervals). Measures of air or water temperature are the most common examples of interval measures. Differences between degrees on a Fahrenheit thermometer represent equal intervals, but only represent "relative" proportional magnitudes. Zero degrees Fahrenheit does not represent an absence of heat.

Thus we cannot say 30° is twice as hot as 15° using this common scale of temperature.

RATIO-LEVEL MEASUREMENT

Quantitative variables that have all the properties of interval-level data (rank ordering and equal intervals between numbers) and also an inherently defined zero point are ratio level variables. Proportional magnitudes of ratio data are meaningful as values that satisfy all the properties of the real number system (e.g., heights of things measured from some zero datum, such as depth of water, because 2 feet is twice as high as 1 foot). Counts of equal units (such as dollars or pounds) may also be considered ratio measures provided there is an absolute zero point at which no units being measured exist.

LEVEL OF MEASUREMENT AND APPROPRIATE STATISTICAL ANALYSES

Appropriate statistics for data measured at a higher level of measurement are often inappropriate for data measured at a lower level. For example, the median (defined below) is an appropriate statistic to use with ordinal level data and can also be used legitimately with interval or ratio level variables. However, it is not an appropriate statistic to use with nominal data, such as a list of names of rivers. No one name in such a list could be called a meaningful "median" value, because the names can not be arranged in any numerical order.

SELECTION BIAS AND DIFFERENTIAL WEIGHTING

Selection bias is a potential data problem for which adjustments may be needed during statistical analysis.

Selection bias occurs when a particular group of subjects is over- or under-represented in the sample. Ideally, the researcher should have planned the sampling strategy carefully enough to avoid this problem as it may compromise the results of the analysis. For example, a household sample of a population with 40% Hispanic households that resulted in only 20% of the sample being Hispanic would under-represent the Hispanic group. If it is anticipated that selection bias may be a problem for a particular population characteristic, such as ethnicity, the sample can be stratified on that variable. This will ensure that resulting sample percentages will match the population percentage breakdown for that variable.

Selection bias usually results from sampling or response problems. However, the analyst may decide to intentionally over-represent one sub-group of a population. For example, the opinions of those in all income levels may be desired where subjects in upper income levels are known to comprise a small portion of the target population. To ensure sufficient response from upper income respondents for analysis, it may be necessary to take a larger sampling fraction from the higher income group than would be indicated for a proportionate sample of the population stratified on the basis of income. After sampling the higher income group, say, at three times the rate as the other subjects, the data could be adjusted during analysis to avoid having the wealthy subjects carry three times their proper weight in the sample. It is necessary to assign weights to either decrease the overall representation of over-sampled elements or increase the representation of those under-sampled. Thus, if the upper income subjects were over-ampled by a 3:1 ratio, their weights in analysis of the total sample should be one-third that of the other elements (i.e. the inverse of the sampling fraction by which the elements were selected).

Another situation in which weighting of cases may be necessary is when selection bias occurs unintentionally and is detected only after the data have been collected. Unintentional bias usually occurs in surveys because of disproportionate amounts of non-response from certain portions of the sampled population. The best way to detect this selection bias is by comparison to known data on the population from which the sample of survey respondents is drawn. Suppose the researcher knows from previous studies, or from other sources such as the Census Bureau, that the general population contains about 50% males and 50% females. Yet, in a general population survey, the sample collected yields 75% males and 25% females. The survey process has resulted in females being under-represented and males over-represented. The appropriate adjustment is to weight the sample according to the known proportions in the target population. To continue this example, suppose the average height of males in the sample is 72 inches and that of females is 65. The adjusted estimate of population average height is 1/2(72) + 1/2(65) = 68.5 inches. The unadjusted estimate is (3/4)72 + 1/4(65) = 70.25, which is obviously biased towards the taller males. So, if the researcher knows the actual proportions of representation in the population, each case should be weighted with that known proportion.

Two important final points about weighting should be kept in mind by the analyst. The first is that biases appearing for variables measured in a survey sample may not have to be corrected by weighting, provided those variables affected are not of interest to the analyst and have no relationship to the results of the study. The second point to remember is that weighting should never be considered a substitute for a well designed survey sampling strategy. This is because only the biases which can be detected are correctable by a weighting procedure. Data resulting from a poorly designed sample are likely to contain many undetectable biases which can seriously affect the study results.

MISSING DATA

After the data have been collected, the researcher will often find a number of incomplete questionnaires. Respondents may have left responses blank because they failed to understand the question or they simply did not wish to answer it. If a particular item is consistently left blank on a large proportion of surveys, the analyst should be alerted that there may be a problem with that item. Prior to making a decision as to how to handle missing data, it is very important to ascertain the reason the data are missing. There are various ways of handling missing responses. We will discuss three: 1) reporting the omissions in special categories, 2) deleting the missing data observations from the analysis, or 3) assigning values to the missing data on the basis of related information.

REPORTING MISSING DATA CASES

The simplest method for dealing with missing data is simply to report them in a special category such as "not ascertained" or "missing", as illustrated below.

TABLE 2

Method 1: Reporting Omissions

Responses	Frequencies	Percentage
< $10,000	15	15
10,001-30,000	45	45
> $30,000	30	30
Missing	10	10
Total	100	100

DELETING MISSING DATA CASES

A second method is to simply delete the cases which have missing data, assuming that the missing data are randomly distributed over the entire sample. The remaining data are then presented as representing the entire sample. For example, suppose an item has three possible response categories and the survey produces 10 non-responses out of a total sample of 100. Only the 90 responses would be tabulated, essentially ignoring the 10 non-responses. However, in the written explanation of the tabulation of a particular item using this approach, the analyst should also disclose the percentage of non-response. Method 2 below shows the data reported with the 10 non-responses deleted.

TABLE 3

Method 2: Deleting Missing Cases

Responses	Frequencies	Percentage
<10,000	15	17
10,001-30,000	45	50
>30,000	30	33
Total	90	100

ASSIGNING VALUES TO MISSING DATA CASES

The third method, assigning values to missing data cases, requires more work. This method is usually used only when the data are critical to the analysis, such as when the item is to be combined with other items in the survey to form an index or scale, or if the analysis requires the use of multivariate techniques (analyzing two or more variables together). In most univariate (single variable) analyses missing cases will not be as critical.

In assigning values to missing data, information about other characteristics of the particular respondent is used to estimate what would likely have been the response to an unanswered question. An example would be a married female respondent who has failed to answer whether or not she works outside her home. The first step in determining her likely answer is to examine her answers to other questions in the survey which describe characteristics helpful in making a decision as to whether this subject works. This is done by looking at variables such as her marital status and the age(s) of any children. The second step is to use known cases to calculate the percentage of working and non-working wives for different child age groups. The result will be a contingency table, as shown below, which estimates the probability that a wife works.

TABLE 4

Method 3: Contingent Probability Estimation

Group	% Working	% Not Working	Total
No children	75	25	100
Preschool children	20	80	100
School age children	50	50	100

If the female respondent in the example is married and has a 3 year old child, she corresponds to the contingency table group for which 20% in this category are working and 80% are not. The decision of whether or not to classify our respondent as working is made by generating a random number between 00 and 99. If this number falls in the category 00-19 the woman is coded as working; if it falls in the category 20-99 she is coded as not working. Her data would then be adjusted with this response and the item included in the analyses as any other. This assignment process could involve looking at more characteristics than the two selected in this example.

BASIC STATISTICAL ANALYSIS PROCEDURES

There are two major categories of statistical analyses: **parametric** and **nonparametric** techniques. Parametric techniques are the most commonly used and are probably most familiar to the reader. These techniques include t-tests, Pearson correlation analysis, analysis of variance, and regression. These parametric techniques are more powerful than nonparametric techniques, but are more restrictive in their requirements as to level of measurement of the data and assumptions concerning the underlying distribution of the population from which the sample is drawn. Non-parametric techniques, often referred to as distribution-free statistics, are suitable for situations where the data do not satisfy the requirements for parametric techniques. They include Spearman rank-order correlation, the Wilcoxon Mann-Whitney U test and the Kruskall-Wallis test. The discussion of analysis techniques presented in this manual focuses primarily on the parametric techniques.[1]

---

[1]For more detail on nonparametric techniques, the reader is referred to texts such as Nonparametric Methods for Quantitative Analysis by Gibbons (1976).

Statistical analysis techniques can also be classified by the number of variables simultaneously analyzed by each technique. Basic statistical analysis procedures included in most studies include both **univariate** (one variable at a time) and **bivariate** (two variables at a time) analyses. More complex statistical analysis techniques can be used to analyze several variables simultaneously and are called **multivariate** procedures.

DESCRIPTIVE UNIVARIATE STATISTICS

The first step in any statistical analysis is to compute descriptive univariate statistics for each of the study variables. This often provides enough information to answer the analyst's questions. Univariate descriptive statistics are often sufficient for situations where the analyst is interested only in describing a survey population with the variables included in the survey questionnaire. If no subgroups need to be compared nor relationships among variables examined, the analysis concludes here. However, even for the most complex statistical analysis, generating descriptive statistics for survey variables is still a necessary first step for subsequent data analysis.

Figure 1 is a tree diagram for determining which types of univariate statistics are appropriate for survey data, depending upon the level of measurement of the survey questions. It allows the analyst, through a series of decision choices, to arrive at the most appropriate technique for the desired analysis. The most commonly used univariate statistics are included in this tree diagram and are discussed below.

Data Type:	Level of Measurement:	Appropriate Statistics:
Univariate (1 Variable)	Nominal	Descriptive: Frequency Distribution of Question Responses.
		Central Tendency: Mode.
		Dispersion: Relative Frequency of the Modal Class.
	Ordinal	Descriptive: Frequency Distribution of Question Responses.
		Central Tendency: a. Mode. b. Median.
		Dispersion: a. Relative Frequency of the Modal Class b. Percentiles. c. Range.
	Interval or Ratio	Descriptive: a. Frequency Distribution. b. Grouped Frequency Distribution.
		Central Tendency: a. Mode. b. Median. c. Mean.
		Dispersion: a. Relative Frequency of the Modal Class. b. Percentiles. c. Range. d. Standard Deviation.
		Normality: a. Skewness. b. Kurtosis. c. Lilliefors, or Shapiro-Wilks

FIGURE 1. Univariate Statistics for Survey Data

Descriptive. Descriptive statistics are based on the frequency distribution of question responses. The most commonly used descriptive is a simple tabulation of the number and corresponding percentage of respondents selecting each answer to a survey question. The responses to different answer categories can then be grouped in different ways. This kind of description of survey question results can be done with all types of data, regardless of the level of measurement.

Descriptive for all levels of data involve displaying the profile of percentage response to a survey question. The data below are hypothetical responses of 200 respondents to a question from topical section A of the Compendium (item 67 on page 18) asking attitudes toward moving to a flood-free zone (assuming 0 non-responses).

Attitude Toward Moving:	Percent	Number
1. "Strongly Opposed"	40 %	(80)
2 "Opposed"	25 %	(50)
3. "Neutral"	15 %	(30)
4. "Mildly Approve"	10 %	(20)
5. "Strongly Approve"	10 %	(20)
Total	100 %	(200)

For meaningful presentation of frequencies for survey questions measured on an interval or ratio scale, the researcher should re-code or group the data into distinct exhaustive non-overlapping categories or ranges. For example, question #4 on the first questionnaire page in Section A of the Compendium asks, "How old is your residence (in years)?______" This question would produce interval level data, and answers could range from 0 years to 100 or more years. A frequency distribution of 100 different answers given could require 100 lines to present. It would be more meaningful to group answers and present a distribution of the grouped results, such as shown below:

Group:	Residence Age:	Percent	Number
1.	1 to 5 yrs.	25 %	(50)
2.	6 to 10 yrs.	30 %	(60)
3.	11 to 15 yrs.	20 %	(40)
4.	16 to 20 yrs.	15 %	(20)
5.	Over 20 yrs.	10 %	(10)
	Total	100 %	(180)

Central Tendency. A measure of central tendency refers to a single value which is most representative of all the data points or responses in the sample for a survey question. Univariate statistics include three commonly used measures of central tendency: the mode, median, and mean.

The **mode** is defined as the most frequently observed value or response to a survey question. It is appropriate for all types of data, but is most often used to describe variables of nominal level measurement.

The **median** is the middle value of the measurements when they are ordered in an array (from small to large). It is the value such that 50% of the values are below and 50% are above it. It is also called the 50th percentile. The **median** provides a better measure of central tendency than the mean if the distribution of the data is skewed (see Dispersion below) or contains extreme values. This statistic is appropriate for use with data measured at the ordinal level or higher (interval and ratio data).

The **mean** is the arithmetic average of the values of a variable measured at the interval or ratio level. It is sometimes also reported for ordinal level data, when it can be argued that there "appear" to be relatively equal intervals between the numbered response codes (e.g. the distribution of "attitudes toward moving" numbered 1 to 5 on page 76). The mean is computed as the sum of the values divided by the number of cases. The mean can be dramatically affected by very large or very small values. In the presence of such extreme values (outliers), the analyst should consider using one of the versions of the trimmed mean. These are known as generalized maximum likelihood estimators, such as Tukey's biweight or Andrew's M-estimator. The formula for calculating the mean is: $\overline{X} = (X_i)/n$, where X_i = the i-th case in a sample of size n, for i=1,2,...n.

All summary statistics such as measures of central tendency and dispersion should be calculated prior to grouping the data. If the un-grouped data are unavailable, estimates of the above measures of central tendency may be calculated from formulae found in texts such as Huntsberger, Billingsley and Croft (1975).

Dispersion. Besides reporting a measure of central tendency and describing the response distribution, the researcher should also report measures of dispersion (i.e., how much variability or scatter is present in the data), and, for interval or ratio data, measures of the normality of the distribution of response. Dispersion statistics describe the shape of the distribution of the data. **Percentiles**, the **range**, the **variance**, and its square root the **standard deviation**, are such statistics.

Dispersion may be described, for all levels of data, by the relative frequency of the modal class (the class of grouped data containing the most frequently occurring value). Question responses which can be arranged in numerical order (ordinal, interval, or ratio data) can also be described in terms of percentiles, indicating the percentage of response at or below certain answers. **Percentiles** are commonly reported descriptive of response distributions for ordinal, interval, or ratio data. The p-th percentile is that value such that p % of the measurements are less than or equal to that value if all the observations are arranged in an ascending array (in order of magnitude). The 25th, 50th and 75th percentiles are known as the lower quartile, middle quartile (median), and upper quartile, respectively.

The **range** is simply the difference between the largest and smallest measurement in a sample of ordinal or higher level data. The **interquartile range** is the difference between the lower and upper quartiles and is used as a measure of dispersion for ordinal level or higher data.

The **variance** is defined for interval or ratio variables as the average of the squared deviations: $\Sigma(X_i-\overline{X})^2/(n-1)$ where $\overline{X}$ = the sample mean. The **standard deviation** is the positive square root of the variance. The standard deviation is generally reported, rather than the variance, since the standard deviation is expressed in the same units of measurement as the original data and is more easily interpreted. The variance and standard deviation are approp iate for interval or ratio level data.

<u>Normality.</u> **Normality** statistics help answer such questions as do the data follow a standard normal distribution, a requirement for most parametric techniques. The Shapiro-Wilks and Lilliefors tests, as well as normal probability plots, should be used to test for normality and are available in standard statistical packages such as SPSS and SAS. **Skewness** and **kurtosis** measures are statistics used to test for the normality of a distribution of interval or ratio data. Skewness is a characteristic of the distribution of the data which may be described in terms of departure from the bell-shaped normal curve. The difference between a skewed and a normal distribution is illustrated in Figure 2, again using the residence age example of five relatively equal age groups numbered 1 to 5. The original example presented in the left hand side of the box portrays a skewed distribution. Skewness measures whether the data tend to fall more in one tail of the distribution than the other. The original skewed example on the left is skewed to the right. In contrast, the example on the right shows what a symmetric distribution of the same data would look like, approximating a normal bell-shaped curve. The symmetric distribution has the largest number of observations in the middle, with decreasing numbers on the right and left sides (tails) of the curve.

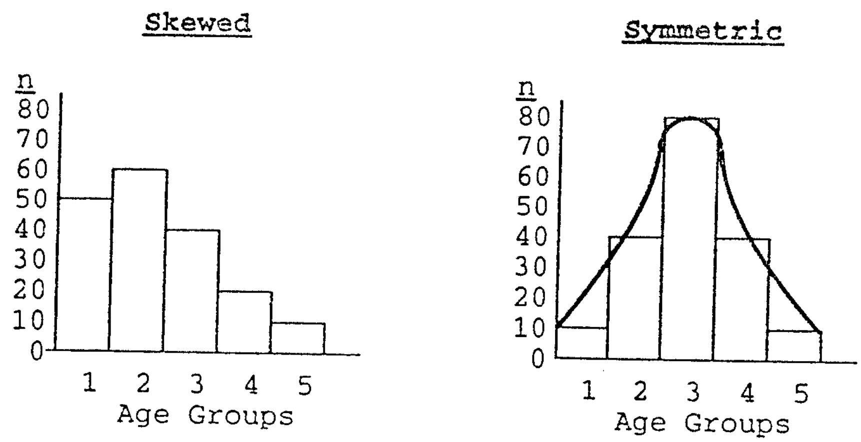

FIGURE 2. Skewed vs. Symmetric Bell-Shaped Distributions

Kurtosis is a measure of whether the distribution is flat (widely spread) or piled up (peaked or grouped tightly) in the center. In general, if a distribution of values is close to normal, the skewness and kurtosis measures will be close to zero.

BIVARIATE RELATIONAL ANALYSIS

Certain statistical techniques focus on examining the relationship or association between two variables. In general terms, the null hypothesis for the two variables X and Y could be H_o: X and Y are statistically independent versus H_a: There exists some association/relationship between X and Y. As previously discussed, the type of relational analysis technique that may be used depends upon the level of measurement of the variables, determined by the types of questions included in the survey questionnaire and how responses are quantified. Different types of bivariate procedures and statistical tests are summarized in Figure 3, and those most commonly used are discussed below.

Data Type:	Level of Measurement:	Appropriate Procedures and Relational Tests for Significant Associations:
Bivariate (2 Variables)	Both Nominal	Association: Chi-Square Test.
	1 Nominal 1 Ordinal	Association: Chi-Square Test. Equality of Group Means: a) Mann-Whitney-Wilcoxon Test (2 Independent Groups). b) Kruskal-Wallis Test (k Independent Groups).
	Both Ordinal	Association: a) Chi-Square (Significance). b) Kendall's Tau (Magnitude). Equality of Group Means: a) Mann-Whitney-Wilcoxon Test. b) Spearman's Rho (2 Indep Grps). c) Kendall's Tau (If Paired).
	1 Nominal 1 Interval (or Ratio)	Equality of Group Means: a) T-test (2 Groups). b) One-Way Analysis of Variance (3 or More Groups). c) Kruskal-Wallis Test (k Independent Groups). d) Mann-Whitney-Wilcoxon Test.
	1 Ordinal 1 Interval (or Ratio)	Equality of Group Means: a) One-Way Analysis of Variance (Interval Variable is Dependent). b) Kruskal-Wallis Test (k Independent Groups). Association: a) Spearman's Rho. b) Kendall's Tau (If Paired). Homogeneity of Variance: F-test.
	Both Interval (or Ratio)	Association: a) Pearson Correlation. b) Simple Linear Regression (1 Independent & 1 Dependent).

FIGURE 3. Bivariate Relational Tests and Statistics for Survey Data

<u>Contingency Table Analysis.</u> The technique most often used for bivariate relational analyses involving two nominal level variables is called **contingency table analysis**. The statistical test most often used to test for significant relationships between variables in contingency table analysis is the **Chi-square test**. Below is an example of a relational analysis using the Chi-square test for hypothetical data from 400 respondents to questions taken from the boating questionnaire in Section N of the Compendium (questions 12 and 14 on page 3). The "n" value in each of the six cells in the contingency table are the number of survey respondents who answered both questions with each of the six possible paired answer combinations. For example, of the 140 respondents who keep their boat at home (column 1), 20 use their boat year around (row 1), and 120 (row 2) do not. The "e" values are the "expected" numbers of answer combinations, shown below.

Where do you keep your boat during the off season?

		AT HOME	PRIVATE MARINA	PUBLIC MARINA	
Do you use your boat year around?	YES	n=20 (e=70)	n=60 (e=60)	n=120 e=(70)	n=200
	NO	n=120 (e=70)	n=60 (e=60)	n=20 (e=70)	n=200
		n=140	n=120	n=140	

Chi Square =142.86
Degrees of Freedom=2
Probability= p<.001

Total Sample = 400

The hypothesis for the above contingency table analysis could be:

H_a: During the off-season those who use their boat year around are most likely to keep it at a public marina, and those who do not use their boat year around are most likely to keep it at home; whereas equal proportions of those who use their boat year around and those who do not are likely to keep it at a private marina. The null hypothesis would be H_O: There is no difference between the distributions of those who use their boat year around and those who do not in terms of where they tend to keep their boat. That is, the null hypothesis can be stated as: whether people use their boat year around and where they tend to keep it are statistically independent variables.

The Chi-square value is calculated by a formula which compares the "n" value in each cell resulting from a survey (e.g. 20 in the Yes/At Home cell of the above contingency table) to the value which would be "expected" if there were no relationship between the two variables of interest. For this purpose a hypothetical "expected value" is calculated for each cell in the table (e.g. 70 in the Yes/At Home cell of the above contingency table). Expected values for each of the six contingency table cells were calculated by multiplying the row "n" (i.e. 200 for the "yes" row) by the column "n" (i.e. 140 for the "At Home" column) and dividing the result by the total sample "n" (i.e. 400). Two assumptions with regard to these "expected" values must be satisfied for the results of Chi-square analysis to be valid. No more than 20% of the "expected" cell values should be less than 5.0 and no individual "expected" value should be less than 1.0.

Individual cell values are calculated for each cell of the contingency table and then summed to derive the Chi-square value. The calculation for each cell value is done by squaring the difference between the observed and expected values and dividing this result by the expected value. A Chi-square value of 142.86 was obtained for the above contingency table example $[2((20-70)^2/70)+2((120-70)^2/70)+2((60-60)^2/60)]$. It is large enough to indicate that the null hypothesis H_0 should be rejected at $p < .001$. The p-value for this analysis is obtained by comparing the calculated Chi-square to those in a table of Chi-square values (found in most statistics books) for the appropriate degrees of freedom. There are two degrees of freedom for this example, calculated as the number of rows minus one (2-1=1) times the number of columns minus one (3-1=2). Since the p-value is less than .05, the test result is statistically significant and the null hypothesis is rejected; it can be concluded that there is a statistically significant relationship between the two variables as specified in H_a. In this example, how the subjects responded to the 'When used boat' question is related to how they responded to the 'Where kept boat off-season' question in some manner.

The nature of this relationship is determined by examination of the data in the contingency table. Sometimes the analyst is surprised by a resulting statistically significant relationship, in which the independent and dependent variables turn out to be related to one another in a manner different from the original H_a hypothesis. If so, the resulting relationship should be reported for what it is, and the analyst should attempt to identify a logical explanation for it.

Correlation Analysis. Another bivariate relational analysis technique is **Pearson's Product Moment** correlation analysis. This technique measures the degree of "linearity" in the relationship between two variables measured at the interval or ratio level. The degree of linear correlation between two variables is measured by the correlation coefficient "r", where "r" represents the sample estimate of the population correlation coefficient "ρ" rho. The correlation coefficient "r" is between -1.00 and +1.00, where plus or minus 1.00 represents a perfect correlation. Values of r close to -1.00 imply a strong negative correlation between the two variables, whereas r close to 1.00 implies a strong positive correlation. A negative correlation means that as variable X increases, variable Y tends to decrease; a positive correlation means that variables X and Y tend to increase or decrease together. When r is close to zero (.00), there is little or no linear association between the two variables, although such variables may be related to one another in a non-linear fashion.

Two questions from topical section A of the Compendium which could be used for a correlation analysis are: Question 18 on page 5 and Question 35 on page 18. They provide data at the appropriate level of measurement (interval) and logically appear to be related to one another. Question 18 asks the market value of home contents and question 35 asks the dollar value amount that flood damages to the home are decreased by measures taken to minimize them. The hypothesis for correlation analysis could be, H_a: Market value of home contents is related to the dollar amount by which potential flood damages to the home are decreased by measures taken to minimize them. The null hypothesis to be statistically tested would be H_o: There is no relationship between market value of home contents and amount of decreased flood damages from measures taken to minimize them. An "r" value such as .85 (with a corresponding low p value such as .01) resulting from computing the correlation for these two variables would indicate rejection of the null hypothesis, and a conclusion that the hypothesized relationship (H_a) holds true.

Each variable for which Pearson Product Moment Correlations are calculated are assumed to be normally distributed. Either **Spearman's Rank** correlation or **Kendall's Rank Order** correlation are the nonparametric tests to use when the normality assumption is violated, when data are measured at the nominal or ordinal scale (less than interval level), or when the sample size is very small (<30 cases).

T-Tests and Analysis of Variance. There are numerous instances in which the analyst is interested in comparing two or more subgroups described by a categorical (nominal level) variable(s) in terms of their group means of a continuous (interval or ratio level) variable. The subgroup variable(s) is the independent variable and the continuous variable is the dependent variable. The subgroups compared can also be termed separate "populations" within the overall population being studied, for example water vs. non-water forms of shipping. The analyst might be interested in whether loading time in minutes for a certain quantity of a given commodity (dependent variable), differs for waterway shipping as opposed to non-waterway forms of shipping (independent variable). The hypothesis might be H_a: The average loading time for a given commodity for shipping by water is less than that for non-water forms of shipping. This would be expressed in the null hypothesis H_o: There is no difference in loading time for shipping a given commodity by water as compared to non-water forms of shipping.

<u>**a. T-Tests**</u>: The appropriate analysis required in the situation where there are two groups (populations) to be compared on the basis of their mean on an interval or ratio level variable is the **t-test**. The assumptions required for the t-test are: 1) independent random samples from each population, 2) data for each population are normally distributed, and 3) populations have common variance. With respect to the second assumption, slight departures from normality may be tolerated. A cursory examination of the skewness and kurtosis measures and a histogram of the data are satisfactory to determine whether the data approximately follows a normal bell-shaped curve (e.g. see Figure 2). With respect to the third assumption, the **F-test** for equality of variance should be performed and is a standard option in all major statistical packages. Most computer programs for the t-test will perform the F-test and will provide two different

"t" values to choose from, one for data which meets the assumption of equality of variances (pooled variance estimate) and one for data which does not meet this assumption (separate variance estimate). The nonparametric equivalent to the t-test is the Mann-Whitney U test which can be used in cases of extreme departures from normality.

The "paired" t-test should be used when the two samples are not independent (assumption #1 above). For example, in before and after treatment experiments, two sets of measurements are taken on the same subjects, the two samples are therefore not independent of one another. If the level of measurement precludes a calculation of the mean as a measure of central tendency and two groups are to be compared on the basis of their medians, a nonparametric equivalent to the t-test should be used: the Mann-Whitney U test (two independent samples) or the Wilcoxon signed rank test (paired samples).

b. One-Way Analysis of Variance: Analysis of variance (**ANOVA**) is an extension of the t-test used to compare means from three or more groups or populations. A t-test can only compare two groups. For example, Question 2 in Part III of the Shipper Interview Form in Section J of the Compendium asks for loading and unloading time of commodities by three modes of shipping: barge, rail, and truck. To test for differences in average loading times between these three modes of shipping it would be necessary to use analysis of variance rather than a t-test. The null hypothesis for this test would be H_0: There is no difference in mean loading time for a given commodity among barge, rail, and truck modes of shipping. The research hypothesis could be H_a: At least one of these three population means differs from the others.

For this example, mode of shipping is the independent variable (or factor variable) and loading time is the dependent variable. The average amount of loading time is hypothesized to depend upon the mode of shipping used. The assumptions required for analysis of variance are basically the same as those for the t-test. A "**one-way**" ANOVA is required to test this hypothesis because only one factor (independent variable) is involved. An n-way ANOVA is defined by the number of factors examined where n=number of factors. In an n-way ANOVA, the analyst is also interested in examining the relationship between the factors.

If the results of a one-way ANOVA are found to be statistically significant and H_0 is rejected (differences exist between two or more group means), then a **multiple comparisons** test can be performed to detect which group means differ. There are a number of multiple comparison tests available depending on how conservative the analyst wants to be. The **LSD** (Least Significant Difference) test is one of the least conservative and would tend to produce the most pairwise significant differences between means for different modes of transport when the differences are not significant. **Scheffe's S** test is one of the more conservative and tends to produce fewer pairwise significant differences when differences do not exist. The recommended multiple comparisons test is **Duncan's New Multiple Range Test.** It is widely accepted and used because it is a compromise between the LSD method and Scheffe's method.

Simple Linear Regression Analysis.[1] Basic regression analysis, with only one independent, or predictor, variable and one dependent variable, is called simple linear regression. It is actually an extension of bivariate correlation analysis, as discussed previously, and it allows the analyst to examine the linear relationship between the dependent and independent variables. The two variables must be measured at interval or ratio levels if regression analysis is to be used.

The equation for the two variable regression model is written and portrayed graphically in Figure 4, where the Y_i's are the observed measures of the individual values of the dependent variable. The $\hat{Y}_i$ values are the estimates of the regression model, portrayed as a straight line, and the X_i's represent the corresponding individual values of the independent variable.

The regression line is the best estimate ("fit" of the line) for the actual observation points scattered around it. The term "a" represents the Y-intercept or constant; it is the value of Y when X=0. The value of B is the slope of the regression line and

---

[1] Simple linear regression is commonly referred to as bivariate regression in research methodology texts (e.g. in Using Multivariate Statistics, by Tabachnick and Fidell, 1989). By contrast, statistical literature commonly discusses the simple linear regression model in terms of the number of predictor variables (e.g. in Gunst and Mason, 1980). In this context, they refer to simple regression, with one predictor variable, as univariate or single variable analysis. The analyst should be aware of these differences in terminology.

is equivalent to the expected change in the dependent variable Y with a one unit change in the independent variable X. The term "e" is the random error in the model. This term, which represents the difference between an actual Y observation and a predicted $\hat{Y}$ observation, is called the residual or error. The underlying principle behind regression is to select the "a" (intercept) and B(slope) values in such a way as to minimize the sum of the squared residuals or error terms, $(Y_i-\hat{Y}_i)^2$. This is why the regression line is called the least squares line.

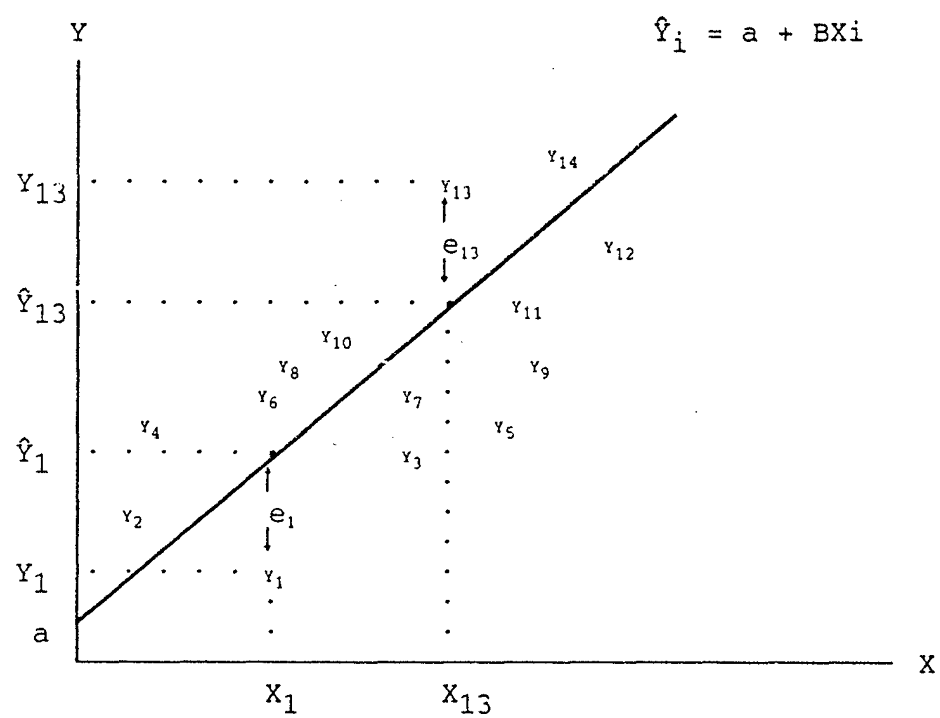

FIGURE 4. Elements of Simple Linear Regression

MULTIVARIATE MODELING PROCEDURES

Occasionally, the survey analyst gathers data for the purpose of developing a multivariate model, which either predicts or describes a relationship between three or more variables. A model is simply a mathematical formulation or equation which captures or depicts the particular phenomenon of interest. The derivation of these statistical models is based on complex parametric techniques such as multivariate analysis of variance, multiple regression analysis or trend analysis. Figure 5 summarizes different multivariate procedures which may be chosen for analysis, based on the level of measurement of the variables to be analyzed. Those procedures most commonly used are discussed below.

Loglinear Models. A loglinear model is a type of multivariate frequency analysis used to test for associations between a set of three or more nominal level variables. One of the variables can be considered the dependent variable and the others the independent variables. Chi-square tests are used to determine which independent variables are significantly associated with the dependent variable and with other independent variables.

Data Type:	Level of Measurement:	Appropriate Multivariate Procedures:
	3+ Nominal Variables	Loglinear Models: Multiple Contingency Table Analysis.
	2+ Nominal Independent Variables, & 1 Interval (or Ratio) Dependent Variable.	N-Way Analysis of Variance: Test for Equality of dependent variable means, for the different groups or categories of the independent variables.
Multivariate (3+ Variables)	1+ Nominal Variable(s) & 2+ Interval (or Ratio) Dependent Variables.	Multivariate Analysis of Variance: Test for Equality of each dependent variable mean, for the different groups or categories of the independent variable(s).
	2+ Interval (or Ratio) Independent Variables, & 1 Interval (or Ratio) Dependent.	Multiple Regression: To build a model to predict, explain, or describe the relationship of the dependent variable to a set of independent variables.
	2+ Interval (or Ratio) Independent Variables, & 1 Nominal (or Ordinal) Dependent Variable.	To build a model to predict, explain, or describe the relationship of a dependent variable that is less than interval level to a set of independent variables. a) Logistic Regression. (Dichotomous Dependent Variable) b) Discriminant Analysis. (Predicts membership of cases in dependent variable groups)

FIGURE 5. Multivariate Statistical Procedures for Survey Data

Analysis of Variance Models. Instead of doing a one-way ANOVA, as illustrated by the example in the previous section, analysis of variance can be expanded to include two or more factors to produce an **N-way ANOVA**. A two-way ANOVA could be done by adding "type of commodity" to the previous example concerning loading times and shipping modes as a second factor. In a two-way ANOVA such as this, it becomes possible to examine the relationship (or lack thereof) between the factors (independent variables) as well as the hypothesized independent/dependent variable relationship.

Multivariate Analysis of Variance (MANOVA) is used when there are two or more interval level dependent variables. The objective of this type of analysis is to determine which of the dependent variables are associated with the independent variable. This may be expanded to an **N-WAY MANOVA**, where the model includes more than one nominal independent variable and more than one interval dependent variable.

Multiple Regression Analysis.[1] Multiple regression allows the analyst to examine the linear relationship between the dependent variable Y, also called the criterion variable, and a set of factors or independent variables X_i, i=1,2...p called predictor variables. The goal is to derive an equation which can

[1] It should be noted that multiple regression and other multivariate analysis techniques are very complex and require considerable expertise as well as the appropriate statistical software package. Complicated multivariate techniques should only be performed using standard statistical packages such as SPSS or SAS, both available in PC versions.

be used as a description of this relationship or to predict values of Y as a linear combination of the X_i's. The dependent variable must be measured on at least an interval scale[1]. The independent variables should be measured on an interval or higher level; however, this technique does allow for categorical variables to be incorporated into the model in the form of dummy or indicator variables. A "dummy" variable may be created by coding each of n-1 categories of response for a categorical variable as either zero or one, and entering each into the model as if they were each continuous independent variables.

Multiple regression allows for the testing of a wide variety of hypotheses. For example, suppose we are interested in predicting the amount that flood damages to the home can be decreased by different prevention measures. In this example, change in flood damages would be the dependent variable for the desired model. One measure of this variable could be obtained with question 35 from Topical Section A of the Compendium which asks the dollar value amount that flood damages to the home are decreased by measures taken to minimize them. The first step in the analysis is to decide what measures and other factors (independent variables) could result in some decrease in flood damages. Some variables which might have a relationship to reduction of flood damages are the 15 types of actions listed in

[1] An extension of regression analysis that allows for the analysis of a dichotomous dependent variable (one which takes on only one of two possible values such as 0 or 1) is called logistic regression.

question 31 of this same questionnaire. The null hypothesis for a multiple regression analysis including all of these actions as independent variables would be H_0: There is no linear relationship between decreased flood damages and the set of independent variables specified above. The alternative hypothesis states that a linear relationship exists between the dependent variable and at least one of the independent variables.

Some **assumptions** required for a valid multiple regression are:

1) The form of the model is specified correctly i.e., we have included all independent variables (X_i) that are important or influential in predicting Y; that we have not included any irrelevant variables and that the relationship between each X_i and Y is linear.

2) The residuals have a mean of zero and are normally distributed.

3) The residuals have equal variance.

4) The errors are independent, not autocorrelated (autocorrelation often occurs with time series data).

5) No exact linear relationship exists between any of the predictors. Violations of this assumption lead to multicollinearity.

Steps in Conducting Multiple Regression Analysis. The **first step** is an exploratory analysis of each independent variable and its relationship to the dependent variable. Descriptive statistics should be computed as well as plots of each independent variable with the dependent variable and a correlation matrix of all variables considered for inclusion in the model. The researcher must also base the analysis on a solid theoretical foundation. The testing of specification errors, that all relevant variables are included in the model, is difficult and can only be accomplished with adequate information about the dependent variable and what factors influence this variable.

The **second step** is variable transformation, if needed. As stated in assumption 1), each predictor must be linearly related to Y. If this is not the case, a relevant independent variable may be transformed so that it does bear a linear relation to Y. Such transformations may include taking the natural log of a variable, raising it to a higher power or taking the inverse of the variable. The particular transformation required depends on the functional form of the original relationship of X_i to Y and can be determined by examining plots of each X_i with Y. Consult a standard text on regression analysis for the proper procedure to employ in variable transformations.

The **third step** is to run the regression, review the resulting statistics and test the assumptions. SPSS and SAS produce thorough reports containing all the important statistics which should be reviewed to evaluate the model. These include p-values, indicating statistical significance of the β coefficients for each independent variable in the model, and R^2 (**coefficient of determination**). R^2 represents the proportion of the variance in Y explained or accounted for by the X_i's. No regression model should ever be evaluated solely on the basis of the R^2 value. One may have a very high R^2 and an invalid model due to violations of the assumptions discussed above. Also, adding additional predictors (independent variables) to the model will always increase the R^2 value, though not necessarily improve the model. In addition to producing the raw R^2 value, SPSS and SAS will produce an adjusted R^2 value to account for the effect of additional predictor variables in the model. Finally, each of the five assumptions specified above must be tested. Violations can lead to very serious modeling errors. Again, both SPSS and SAS allow the analyst to request information which can be used to test each of these assumptions.

The first assumption, **correctly specifying the model**, is initially addressed by the analyst's judgement of which independent variables to include. In the example above, the 15 independent variables taken from the flood damage questionnaire in Section A of the Compendium appear to include most of the actions that can be taken to reduce flood damages to contents of residences. After running the regression program for the first time, those actions which are not significantly related to the dependent variable (reduction in flood damages) can be eliminated or transformed in an attempt to better specify the model. Any of the stepwise methods of including variables provided by SPSS or SAS will automatically exclude those variables which are not significantly related to Y.

The **examination of residuals** is crucial in the testing of assumptions two and three. This information should always be requested as output from the regression analysis procedure. A normal probability plot or a histogram of standardized residuals will test the assumption of normally distributed errors. Slight departures from normality are tolerable. A plot of residuals with predicted Y values should be produced to test the assumption of **homoscedasticity** (assumption 3). The plot should reveal a random scattering of points about a horizontal line through zero. The presence of any sort of pattern in the plot of the residuals with the predicted Y values (e.g. larger residuals at one end of the line) indicates unequal variance of error terms. This indicates violation of the assumption and is called **heteroscedasticity**.

Examining standardized residuals and various distance measures will also alert the researcher to the presence of **outliers**. An outlier is an extreme data point (much higher/lower than the others) which, depending on its magnitude, can dramatically affect the computation of the regression line. Each extreme data point must be examined to determine if it is a valid point, a recording error, or some other kind of error. Sound theoretical knowledge about the phenomenon being investigated is necessary to make decisions as to how to handle these influential data points. Outliers should never simply be deleted without careful consideration. The following statistics provide information about the potential influence of a particular data point: **leverage values**, **Cook's distance measures**, **standardized difference in fit**, and the **covariance ratios**[1].

[1] Indicators of influential points are: Cook's Distance>1; leverage values > $2*p/n$ where p=# independent variables, n=sample size; standardized difference in Beta values > $2/n^{1/2}$; covariance ratios CR such that $|CR-1|>3+p/n$.

Multicollinearity results from a high degree of association among two or more independent variables and can have grave consequences (assumption 5). One may even find estimates of the beta coefficients (β_i) to be of the wrong sign due to the inflation of the variances of the beta terms. To test for multicollinearity, a correlation matrix should be produced, displaying all independent variables to be included in the multiple regression analysis. When any two independent variables are found to be strongly correlated ($r = .90$ or greater), the inclusion of both may present problems and the situation should be further examined. Another sign that multicollinearity is present is when the overall F-test is significant but none of the individual regression coefficients are significant. In addition, **tolerance values** should be examined. Tolerance is the proportion of variation in a predictor variable that is independent of the other predictor variables. Small tolerance values (near .01) indicate that the information provided by that predictor is provided by other predictors and is largely redundant. Other statistics which should be reviewed are the **eigenvalues**, **variance decomposition proportions**, **variance inflation factors**, and **condition indices**[2]. The definitions of these terms are somewhat technical and require knowledge of matrix algebra. The reader is urged to consult a statistics text

[2] Indicators of multicollinearity are: condition indices >30; tolerance values near .01; and relatively high variance inflation factors. Variance decomposition proportions close to 1.0 with corresponding eigenvalues near zero point out the variables involved in the multicollinearity.

on regression analysis for a discussion of these terms. Those not experienced in multiple regression analysis should also seek help from statisticians or others with experience using regression analysis.

To develop the **final form of the model**, a regression analysis will usually be performed in an iterative fashion. After examination of the initial model and testing the regression assumptions, the researcher must make decisions about the choice and form of the variables in the model and cases to be included in the analysis. It is often necessary to re-run the analysis several times before deriving the final model.

When the model is in final form, it may be used to describe and/or predict. For either purpose, the researcher must be cautious in extrapolating the results of the model beyond the range of the data used to derive the model. For example, in predicting income, if the data used for analysis included only people age 18 to 40, it would be theoretically incorrect to make predictions about the income of people in the 65-year and older category.

Forecasting and Trend Analysis Models. Forecasting models are developed from measurements on variables observed at regular known intervals over a specified period of time. Multiple regression is often used for this. An example is daily inventory levels of a product over one month. The purpose of forecasting analysis is to discover a systematic pattern in the series so that the behavior of the data can be described by a mathematical model. This model may then be used to predict how the series will behave in the future, or to evaluate the effect of an interruption or disturbance in the series. For example, it may be used to determine if instituting a change in a production process has a significant impact on the quantity of inventory. Forecasting techniques may be divided into two major categories: qualitative and quantitative methods.

<u>a. Qualitative Trend Analysis.</u> Qualitative trend analysis methods require inputs that are mainly products of subjective judgement, intuitive thinking and accumulated information, often from a number of experts in the area of study. Qualitative methods, also known as technological methods, can be grouped into exploratory and normative techniques. Exploratory methods such as Delphi, S-curves, analogies and morphological research, begin with the past and present data as their starting point and move toward the future in a heuristic manner, looking at all possible scenarios. Normative methods such as decision matrices, relevance trees and system analysis begin by examining a future condition and work backwards to discover if that future scenario is feasible given the resources and technologies available. Qualitative forecasting will not be treated here. The reader is referred to texts such as <u>Strategic Planning</u> by G.A. Steiner (1979).

<u>b. Quantitative Trend Analysis.</u> Quantitative forecasting may be performed when three conditions exist: 1) historical data are available; 2) the data are quantitative in nature; and, 3) it

can be assumed that at least some aspects of the observed pattern in the data will continue in the future. There are three major stages in building quantitative forecasting models. They are identification, estimation and diagnostic testing.

Identification involves selecting a tentative model type, selecting the number and kinds of parameters involved, and determining how they might be integrated into the model. This stage involves plotting the series over time to detect any upward or downward trend, determining if a data transformation might be needed and whether the series displays any sort of seasonal or cyclic fluctuations.

Estimating is the process of fitting the selected model to the data, estimating its parameters, and testing these parameters for significance. If the parameter estimates are statistically unacceptable in explaining the behavior of the series, the analyst must return to the identification stage.

In **diagnostic testing**, the analyst examines how well the tentative model fits the data by examining plots and statistics describing the residual or error series. The results of this stage determine whether the model is adequate or whether it is necessary to return to stage 1 and try to identify a better model.

Quantitative trend analysis methods may be further divided into causal (regression) and time series techniques.

Causal/Regression Models assume the variable to be forecast exhibits a cause-effect relationship between one or more independent factors. These models are regression models (e.g. multiple regression, logistic regression), and can be used to predict future values of the variable of interest. The regression techniques employed include ordinary least squares, weighted least squares, two-stage least squares, logistic regression, and models designed to accommodate autocorrelated errors (e.g. Prais-Winsten, Cochran-Orcutt and maximum likelihood estimation).

Time series models predict the future solely on the basis of past values of the variable of interest or on past error terms. The forecasting relationship is a function of time only. The purpose of using time series models is to discover the underlying pattern in the historical data and extrapolate it into the future. No attempts are made to explain the behavior of the data in relation to other factors (i.e., the objective is to predict what will happen but not why). These models can be loosely divided into the following categories: smoothing models, Box-Jenkins (ARIMA) models and models for decomposition of cyclic data.

Smoothing models use exponential smoothing to remove the effect of random fluctuations in the series and give more weight to recent observations. There are a variety of smoothing models available. These techniques are relatively simple and are more appropriate for short-term forecasting.

ARIMA models (Autoregressive Integrated Moving Average) attempt to mathematically describe the disturbances or shocks that occur in a time series. ARIMA models combine as many as three types of processes: autoregression, differencing to achieve stationarity, and moving averages. Discussion of the details of these models is beyond the scope of this manual. The reader is referred to standard texts on this subject such as Box and Jenkins (1976) or McCleary and Hay (1980).

Cluster and Factor Analysis Techniques. Cluster analysis and factor analysis are two other multivariate techniques which are beyond the scope of this document. Both include a variety of cluster and factor analysis alternatives. These techniques are most often used for dividing a series of variables into similar clusters or factors, or for grouping survey respondents (subjects or objects) based on their responses to a series of variables. The level of measurement required for variables to be factored or clustered varies, depending upon the particular computerized technique selected. For detailed discussion of these techniques, the reader is referred to authoritative texts such as Harmon's Modern Factor Analysis (1967), and Everitt's Cluster Analysis (1980).

CHAPTER VII

REPORT WRITING

It is rare that a survey receives very much documentation in a study report. This is unfortunate, because as Rea and Parker (1992) have pointed out, the survey process is not really complete unless it is documented. The better the process is documented, the more useful it will be to reviewers, data users, and potential employers of the survey methodology.

CRITICAL ITEMS TO REPORT

There are some items which are always critical to include in a report. Survey and analysis procedures should be fully documented. The survey procedures can be included within the main body of the report, in an appendix, or as a separate report, but should always be reported in some form. The report should also provide a detailed description of the basic survey results. A brief summary, which will not add a great deal of volume to the report, is also something which should generally be added to either the beginning or end of the document.

There are certain other elements of all survey efforts that are also critical to include in the report, regardless of reporting format. These critical elements would include a statement of major study objectives, a description of the type of survey instrument that was employed, and a definition of the sampling frame. These items give reviewers and potential users of the data a greater knowledge of the questions being pursued, confidence that the survey procedures were reliable, and a greater understanding of the results. The reader should be reassured that the analysts were aware of potential biases in the methods used and that efforts were taken to minimize these biases.

ELEMENTS OF A DETAILED REPORT

EXECUTIVE SUMMARY

The executive summary provides an opportunity for those with limited time to review the results and implications of a survey. It is generally found at the very beginning of a report and can range from one to ten pages or more in length. The executive summary may be limited to only a short list of major findings, or it may include an expanded discussion of all of the critical elements described above.

BASIC OUTLINE FOR A DETAILED REPORT

It is helpful to think of the main body of a report as describing the first ten steps of the survey process as discussed in Chapter I of this manual. The introductory chapter usually defines the survey objectives, provides background information, and reviews related literature. A separate chapter is sometimes needed for literature review when an extensive body of literature exists on the subject of the study. A second or third chapter covers the general survey procedures; including survey method selected, design and pretest of the questionnaire, drawing a sample, personnel selection and training, implementation of data collection, assessment of non-response, and preparation of the data for analysis. The remainder of the report describes the analysis, results and implications. There should be enough information provided for all steps so that another analyst could replicate the survey process in the same manner.

INTRODUCTION

The introduction should describe the purpose and scope of the study for which the survey was conducted, and the nature of the specific questions that were investigated. It should describe previous work on the subject and deficiencies with existing information. The introduction should explain why a survey was used to obtain this information, as opposed to an existing data source. This discussion might focus on how existing data might be too old, not localized, or not focused on the specific issues of interest for the study.

This chapter should state and discuss the specific study objectives, together with any hypotheses that were formulated. For example, the objective of determining the extent to which depth of flood waters is a determining factor in damage estimation could be discussed. The chapter could then describe the hypothesis that depth of flooding relative to the first floor of a building is a significant factor in determining the percent damage to structure and contents.

GENERAL DISCUSSION OF SURVEY METHODOLOGY

An early chapter should include discussion of all the steps in the survey research process, from selecting the type of survey method through the data editing process.

Describe the Selection of Survey Method. The report should explain why the particular type of survey vehicle or combination of survey vehicles were employed. This discussion might include how priorities such as response rate, financial constraints, time constraints, and the complexities of the questions affected the decision.

Recount How Questionnaire Was Designed. The report should describe how the questionnaire was designed and why the particular types of questions were used. Reference should be made to the appropriate section(s) consulted from the Compendium of approved survey questions. This section could explain why a particular style of wording may have been used and anything that may have been done to reduce bias in the questionnaire or to better communicate with the people in the sample frame.

Review the Pre-test Procedures. A description of the pre-test procedures should include a discussion of how the pre-test respondents were selected, the number of respondents involved in the pre-test, the types of problems encountered during the pre-test, and the types of changes that resulted in the wording of the questionnaire or the survey process.

Describe the Sampling Frame. The sampling frame refers to the specific population from which the sample is drawn. It may be the entire population of a geographic area or it may be more narrowly defined by eliminating one or more specific population units or subgroups. The report should explain why a particular sampling frame was selected and whether or not it is representative of a larger group. When possible, the discussion should also include a description of the general demographic characteristics of the population being surveyed. These are data that should be available from the census for population surveys, and can be corroborated by survey demographic questions after the data are collected.

Describe Sampling Procedure. The methodological issue which is generally of most concern is the sampling procedure used in the survey. One of the first concerns of any reviewer or potential user of the survey results is whether the sample was representative of the population for which generalizations are to be made. The report should describe the specific sampling strategy employed and the procedures followed to select individual respondents. There should be a reporting of the number of respondents contacted and the number for whom the survey was successfully completed. It is important to report this completion rate, including, when appropriate, a demographic breakdown of the sample characteristics compared to the population characteristics.

Describe the Personnel Selection and Training Procedures. The report should discuss the source of personnel used in the survey process and what steps were taken to prepare them for the survey process.

Discuss the Data Collection Process. This section should include a discussion of scheduling, supervision, and any significant logistical concerns for carrying out the survey.

For a mail survey, this section must describe the mail-out package(s) of materials used and the cover letters accompanying each mailing.

Describe the Assessment of Non-response. If the number refusing to participate in the survey is considered to be significant, then the reasons for the non-responses should be assessed and listed in the report, along with a reporting of any systematic bias that might result from the non-response. As stated above, the report should describe any steps that were taken to compare the demographics of the selected respondents who did not respond to the survey to those that did.

Discuss Procedures for Coding, Entering, and Editing Data. A short discussion of the data entry and editing procedure could be useful to other analysts contemplating the same type of survey. The criteria and procedures used for eliminating any faulty survey forms, data errors, or extreme responses should be described.

DESCRIBE DATA ANALYSIS AND REPORT RESULTS

Describe Data Analysis. The methods of data analysis used should also be described in detail. This must include a discussion of all statistical procedures that were used. It is usually unnecessary to identify the specific computer software used in the analysis, unless unique versions of statistical techniques are used which can only be found in a particular software package.

Reporting Results. The results of the study should be presented in appropriate detail with respect to the primary study issues, but this should be reported as concisely as possible. Only response statistics should be reported. Extraneous detail from computer packages should be eliminated from the presentation. It is helpful to use a word processing program to edit tables and a graphics package to edit figures that are output from statistical packages.

DESCRIBE IMPLICATIONS

In describing the survey implications, it should be noted whether the survey results apply only to the current situation or whether they can be extrapolated to other projects, the entire district, region, or the nation. It is also important to indicate whether the results appear to only be relevant to one specific point in time or whether the results might be useable for some time to come. The discussion of implications should, when possible, include comparisons with previous surveys. It may go as far as to include tables comparing the major conclusions to previous research.

APPENDICES

Survey documentation can lend itself to many pages of tables, graphs, and even case-by-case listings of surveys. To limit the size and enhance the readability of the report, it is recommended that voluminous detail that may be of only specialized interest should be relegated to the report's appendices. Care should be taken not to include anything in an appendix which would breach promises of confidentiality or non-disclosure made to survey respondents. For example, unsolicited comments from survey respondents are often included verbatim in an appendix of the final report. Some respondents will sign their name after writing a comment about which they have strong feelings. In transcribing such a comment the name should be omitted.

The appendices should contain copies of the survey forms used. Including the survey questionnaire and other supporting materials can make it substantially easier for other interested parties to replicate the survey procedures. This can save substantial resources on future survey efforts. The survey form itself can be most useful if it is annotated, with a written rationale for each section of the survey, if not individual items. A question-by-question annotation should be considered if the survey is considered particularly important, or if the nature of the questions is somewhat more complicated than usual.

Supporting material may include pre-survey announcement letters and press releases; cover letters, post cards, and follow-up reminders for mailed surveys; letters of introduction, and introductory statements leading into face-to-face surveys; and scripts from telephone interviews. Each exhibit can be accompanied by a paragraph describing its purpose and how it was used.

CHAPTER VIII

SUMMARY AND RECOMMENDATIONS

SUMMARY

This manual was prepared as an instruction guide on use of the Compendium of OMB approved survey questionnaires, for use by Corps economists and other planners responsible for conducting surveys. It builds upon earlier IWR publications in the NED Procedures Manuals Series, giving specific direction on how to identify and adapt appropriate OMB approved questionnaire items from the Compendium. It also gives more detailed explanation and direction than previous manuals on appropriate procedures for the analysis and reporting of survey results.

The first three chapters of this Manual refer directly to the Compendium and how it is to be used. Chapter I provides an introductory background to the Compendium in the context of Corps survey efforts. It ends with an overview of the steps of the survey process. Chapter II describes ways to cross-reference the contents of the OMB Compendium by topic of study, different methods of survey data collection, and different types of survey questions. A cross-classification table provides a quick and easy way to locate desired questionnaire items with respect to these reference criteria. Chapter III provides guidance for adapting Compendium questionnaire items to specific study purposes and methods of survey delivery.

The remaining chapters in this Manual provide more generic information on how to design and conduct surveys. Particular emphasis is placed upon the steps of the survey process not covered in previous manuals, especially data analysis.

In sum, this Manual attempts to achieve several important purposes. It serves as a guide to use of the OMB manual of approved questionnaires. It builds upon and adds to the survey research material found in previous NED manuals, and refers the reader back to these previous manuals and to other references for specific direction and examples.

RECOMMENDATIONS

The Compendium is an ad-hoc collection of survey questions, which has had some additions made at three year intervals, resulting from review of planning needs and other considerations. A more detailed review and revision of the Compendium could improve its applicability to contemporary Corps planning needs, and the reliability and validity of all survey items.

The first step in this revision should be a survey of Corps planners, to identify current survey data needs. Results of the survey may indicate the need for additional topical areas serving, for example, environmental or other new areas of Corps planning emphasis. Survey results may also indicate the need for more examples of certain types of questionnaire design, such as telephone questionnaires, noted in Chapter II as missing from most topical sections.

Every questionnaire item to be retained in, or added to, the Compendium, should then be reviewed for validity and reliability. Where necessary, items should be revised and improved before submitting the Compendium for its next three year OMB approval.

The Compendium should then be redesigned to make it more user friendly. The format should be changed so that it is presented more as a catalogue of survey questions, than the present ad-hoc collection of survey questionnaires. Ultimately, the Compendium format should be transformed into a computerized catalogue of questions, which Corps planners could use to compile survey questionnaires more efficiently and effectively than is presently possible.

LITERATURE CITED

Box, G.E.P. and Jenkins, G.M. 1976. Time Series Analysis, Forecasting, and Control. Revised Edition. San Francisco; Holden-Day.

Dillman, D.A. 1978. Mail and Telephone Surveys: The Total Design Method. New York; John Wiley & Sons.

Everitt, B. 1980. Cluster Analysis. New York; Halsted Press.

Gibbons, J.D. 1976. Nonparametric Methods for Quantitative Analysis. New York; Holt, Rinehart and Winston.

Gunst, R.F. and Mason, R.L. 1980. Regression Analysis and Its Application. New York; Dekker, Inc.

Hansen, W.J.; Mills, A.S.; Stoll, J.R.; Freeman, R.L.; Hankamer, C.D. 1990. "A Case Study Application of Contingent Value Method for Estimating Urban Recreation Use and Benefits." IWR Report 90-R-11. National Economic Development Procedures Manual - Recreation, Volume III. U.S. Army Corps of Engineers Water Resources Support Center, Institute for Water Resources. (November) Ft. Belvoir, Virginia

Harmon, H.H. 1967. Modern Factor Analysis. Chicago; University of Chicago Press.

Hollander, M. and Wolfe, D.A. 1973. Nonparametric Statistical Methods. New York; John Wiley & Sons.

Huntsberger, D.V.; Billingsley, P.; Croft, D.J. 1975. Statistical Inference for Management and Economics. Boston; Allyn and Bacon.

Kish, L. 1965. Survey Sampling. New York; John Wiley & Sons.

Kraemer, H.C. and Thielman S. 1987. How Many Subjects? Statistical Power Analysis in Research. Los Angeles, Sage Publications.

McCleary, R. and Hay, R.A. 1980. Applied Time Series Analysis for the Social Sciences. Beverly Hills, Ca.; Sage Publications.

Mills, A.S.; Davis, S.A.; and Hansen, W.J. 1991. "Primer for Surveying Flood Damage for Residential Structures and Contents." IWR Report 91-R-10. National Economic Development Procedures Manual - Urban Flood Damage, Volume II. U.S. Army Corps of Engineers Water Resources Support Center, Institute for Water Resources. (October) Ft. Belvoir, Virginia

Moser, D.A., and Dunning, C.M. 1986. "A Guide for Using the Contingent Value Methodology in Recreation Studies." IWR Report 86-R-5. National Economic Development Procedures Manual - Recreation, Volume II. U.S. Army Corps of Engineers Water Resources Support Center, Institute for Water Resources. (March) Ft. Belvoir, Virginia

Ott, L. 1977. An Introduction to Statistical Methods and Data Analysis. Belmont, California; Wadsworth Publishing Co.

Rea, L.M., and R.A. Parker 1992. Designing and Conducting Survey Research: A Comprehensive Guide. San Francisco, Jossey-Bass Publishers.

Schaeffer, R.L.; Mendenhall, W. and Ott, L. 1979. Elementary Survey Sampling. North Scituate, Massachusetts; Duxbury Press.

Steiner, G.A. 1979. Strategic Planning. New York; The Free Press.

Survey Sampling, Inc. 1988. The 1988 Unlisted Rates of the Top 100 MSA's. 2 p. Fairfield Connecticut.

Tabachnick, B.G. and Fidell, L.S. 1989. Using Multivariate Statistics. New York; Harper and Row.

The Privacy Act of 1974. Public Law 93-579, 88th Congress, Statute 1896. Effective, September 27, 1975.

U.S. Army Corps of Engineers, Office of the Chief of Engineers, Planning Division - Directorate of Civil Works 1984. Approved Questionnaire Items for Collection of Planning Data. (January) Washington, D.C.

U.S. Army Corps of Engineers, Office of the Chief of Engineers, Planning Division - Directorate of Civil Works 1990. Engineering Regulation #1105-2-100. (December) Washington, D.C.

APPENDIX

GLOSSARY OF COMMON STATISTICAL ANALYSES

PARAMETRIC STATISTICAL TERMS

ANALYSIS	PURPOSE
1. Analysis of Variance (ANOVA):	To compare dependent variable means for three or more subgroups of the surveyed population, using interval or ratio data. The independent or groups variables (X_j) are categorical (nominal or ordinal data).
2. Chi-square:	To examine the relationship between two categorical variables (nominal or ordinal data) to determine whether or not they are statistically independent.
3. Cluster Analysis:	To identify homogenous groups or clusters of items or cases bearing similar characteristics, where group membership and number of groups is unknown, and using any combination of variables (nominal 0/1, interval, or ratio data).
4. Discriminant Analysis:	To categorize cases into homogenous groups (dependent variable) based on one or more multivariate discriminant functions, where group membership is known, and to identify independent variables most important for classifying a case into a particular group, using any combination of nominal (0/1), ordinal, interval, or ratio data.
5. Factor Analysis:	To identify a relatively small number of factors used to represent relationships among sets of interrelated variables, using interval or ratio data. In addition to variables, some factor analysis programs can be used to factor (produce homogeneous groups) of survey respondents (nominal data), based on their responses to survey questions.
6. Logistic Regression:	To estimate the probability that an event will happen (dependent variable), where the dependent variable is dichotomous (0,1 nominal or ordinal data). The independent variables are continuous (interval or ratio), nominal (0/1), or ordinal "dummy" variables.

7. Log-Linear Models: Multivariate contingency table analysis, with one categorical (nominal or ordinal data) dependent variable, and two or more categorical independent variables. Some continuous variables (interval or ratio data) may also be used as co-variates. Chi-square is used to test for statistical associations between variables.

8. Multivariate Analysis of Variance (MANOVA): To compare group subgroup means for each of two or more dependent (Y_i) variables, where interrelationships between Y_i's are also into account. Independent variables (X_j) are categorical (nominal or ordinal data), and dependent are continuous (interval or ratio data).

9. Pearson Correlation: To measure the strength and direction of linear association between two variables using interval or ratio data.

10. Regression: To build a model (equation) to describe or predict the relationship between a dependent variable which is continuous (interval or ratio data) and one or more independent variables which are either continuous (interval or ratio data), nominal (0/1), or ordinal "dummy" variables.

11. Reliability Analysis: Evaluating the reliability (consistency) of a test or scale of measurement, based upon resulting item and item-to-total correlations, reliability (alpha) coefficients, and various other information depending upon the computer software used (interval or ratio data required).

12. T-Test: To compare 2 group means of interval or ratio data.

NON-PARAMETRIC STATISTICAL TERMS

ANALYSIS	PURPOSE
1. Runs Test:	To test whether or not <u>dichotomous</u>, <u>ordinal</u>, or higher level data are randomly distributed with respect to the median.
2. Runs Up & Down Test:	To test whether or not <u>ordinal</u> or higher level data are normally distributed.
3. Chi-square:	Goodness-of-fit test for whether or not <u>categorical</u> data follow a particular probability distribution, such as "uniform".
4. Kolmogorov-Smirnov test:	Goodness-of-fit test for whether or not ordinal or higher level data follow a particular distribution, such as "normal", "poisson", or "uniform".
5. Sign Test:	To test whether or not the median, for <u>dichotomous</u> data, equals a specified value for a one sample problem. For paired observations the median of differences is used. Confidence intervals can be calculated. Can be used when assumptions for the paired T-test are not satisfied.
6. Wilcoxon Signed Rank Test:	To test whether or not the median, for data which are <u>ordinal</u> or higher level and symmetric about the mean, equals a specified value for a one sample problem. For paired observations the median of differences is used. Confidence intervals can be calculated. Can be used when assumptions for the paired T-test are not satisfied.
7. Mann-Whitney U Test:	To compare medians of two independent samples of <u>ordinal</u> or better data. Can be used when assumptions for the paired T-test are not satisfied.
8. Kruskal-Wallis Test:	To compare medians of k independent samples of <u>ordinal</u> or better data. Can be used when assumptions for a One-Way Analysis of Variance are violated.
9. Siegal-Tukey Test:	To test, using <u>ordinal</u> data, whether or not two populations have equal variance, with medians equal and unknown. Can be used when assumptions of the F-test are violated.

10.	Sukhatme, Siegal-Tukey, Mann-Whitney-Wilcoxon Tests:	For use with ordinal data to test whether or not two populations have equal variance, when medians are known or can be estimated from sample data.
11.	Spearman's Rank Order Correlation:	An approximation of Pearson's Product Moment Correlation which can be used to test for an association between ordinal variables when the normality assumption is violated and the ratio of number of cases to number of variable categories is small.
12.	Kendall's Tau:	Basically the same as Spearman's Rank Order Correlation, but more appropriate when there are a large number of cases and a small number of variable categories. Kendall's Tau also usually produces a more reliable p-value, compared to Spearman's Rank Order Correlation.
13.	Cramer's V, Yule's Q, Goodman-Kruskal Tests:	For use with nominal or ordinal data, to assess the strength of association between a row and a column variable in a contingency table.

*U.S. GOVERNMENT PRINTING OFFICE 1993 -342 -662/82404