# TABLE OF CONTENTS

## Section I, INTRODUCTION



## Section II, CHAPTER by CHAPTER COMMENTARY

# Section III; EXERCISES

## The Author

Although Darren George (the first author of the textbook) is writing this manual, reference is often made to “us”. The other half of us, is, of course Paul Mallery, the second author of the textbook. When the term “us” is used, the material that follows is based on consensual thoughts or findings of both authors even though the first author is writing them. The manual is written in the first person and the tone is conversational. In pages that follow, I share some of my experiences using the SPSS book and provide helpful hints on how to use the book, chapter by chapter, in your own teaching.

## Content and Organization of the Manual

This manual contains three primary sections.

1) **The introduction (pp. 2-10)**: The introduction considers a variety of general issues associated with the structure, rationale, creation of, revisions, application, and other concerns associated with the book and its use. Reading this section will provide you with a (hopefully) pleasurable foray into many issues that motivated the creation of the book, determined the content, and dictated the structure we have followed. Many professors don’t even open the teacher’s manual; you, clearly, have done at least that much. I think that the 15 minutes required to read this section will be well worth your time.

2) **Chapter-by-chapter analysis (pp. 11-46)**: This section considers each chapter individually in some detail. Both Paul and I are teachers. We have extensive experience presenting statistical material at an introductory statistics level, in research methods classes, and in more advanced multivariate classes. These pages deal largely with teaching tips for clear presentation of the contents of each chapter. Only material up to Chapter 16 is included in this manual—the typical coverage in an introductory statistics class. Chapter 17 through 28 are not only advanced and more complex, but teaching these techniques is so closely associated with the context of the class or the teaching style of the instructor that it seems unwise to comment extensively on styles of presentation in a manual of this type. Further, presentation of these more-challenging procedures is typically dictated by the need for an immediate application in a particular setting.

3) **Sample exercises (pp. 47-95)**. There are 10 different data sets available on the Web site at **www.abacon.com/george**. There are no exercises in the body of the text; this manual is designed in part to remedy that lack. The exercises occupy the final 49 pages of the manual. It makes use of exactly two of the 10 data sets at **www.abacon.com/george**. Both files are large, are real data, and provide the necessary resources to perform just about any type of data analysis. Both files are also straightforward and intuitive. The helping file (**helping3.sav**) contains data designed to identify factors that influence helping among friends. The divorce file (**divorce.sav**) explores cognitive and interpersonal factors that affect recovery from divorce. The Exercises sec-

tion is particularly useful because it provides a number of statistically significant procedures that would be difficult for a professor to uncover on his or her own. The exercises include chi-square analyses, t-tests, one-, two-, and three-way ANOVAs, tests for curvilinear trends, and multiple regression analyses. Other types of analyses presented in the first 16 chapters don't require much prior thought and are thus not included: For instance, frequencies, descriptives, and correlations can be conducted easily with just about any set of data.

The use of only two data files is, I feel, very functional. In our data entry chapter (Chapter 3) we explain that much of data analysis is easy if the research project has been thoughtfully constructed and if the variables have been intelligently entered. I further explain (to my students, not in the book) that I have spent four years in the large helping file (over 1000 subjects, over 100 variables) and thus can negotiate quickly through any types of analyses because of my extensive involvement with those data. Students, on the other hand, start from scratch. They have never seen the file before and it takes awhile for them to gain a sufficient feel for the data to be comfortable conducting different analyses. Once they learn about the data, however, then the focus is allowed to switch to the understanding of data analysis (rather than trying to figure out the nature and structure of the file). If a student is given assignments that require him or her to switch from file to file to perform different types of analyses, it is not only frustrating to the student, but also results in some mangled write ups (of ANOVA results, for instance) due, in part, to the lack of a thorough grasp of the data and what the study is about.

For textbook examples, a great deal time was devoted to creating the **grades.sav** file. In the first chapter we introduce the data file that is used in 16 of the 28 chapters. Consistent with the sentiments of the previous paragraph, much time is spent in Chapters 1 and 3 to acquaint the student thoroughly with the content of this file. This file effectively demonstrates many of the procedures, however, the **grades.sav** file is much too limited for an extensive set of exercises in this manual. The file has 105 subjects, 17 variables and fits on two pages at the end of chapter 3. Also, this particular file is entirely fictional, created by us to demonstrate textbook procedures. The **helping3.sav** file and **divorce.sav** files provide the necessary resources to conduct any types of analyses, but, as mentioned above, require some learning before they can be utilized effectively.

All problems related to the divorce file are presented first, organized by type of analysis. Then all problems associated with the helping file complete this section. Here is the structure or format we follow in the exercises.

- The type of analysis is identified (e.g., independent-samples t-tests, one-way ANOVAs) and occasionally certain features of the output are also included (e.g., 3-way ANOVA with 1 × main effect, 1 × 2-way interaction, and 1 × 3-way interaction).

- The question addressed in the problem is presented in words: For example, "Do men and women differ significantly on the amount of empathy they experience?"
- The computer printout is reproduced and placed in a simple box that identifies all relevant output from the analysis. The SPSS output will be rendered in simple Courier font, linear format rather than the SPSS 10.0 boxes.
- Occasionally special features of a particular analysis will be commented upon. Discussion of why the results turned out as they did or instructions of how to write results in correct APA statistical format is up to you.

## Why *SPSS for Windows Step by Step* Was Written

The present book was initially written to make the use of statistical software (SPSS in this case) clear and straight-forward for students. Paul and I have both spent enough of our lives in teaching to have seen many students traumatized by the professor who cheerily says "analyze these data in the lab; read the SPSS manuals for specifics of how to accomplish this." The manuals are 3000 pages long and beyond comprehension to almost anyone other than statistical experts. For Paul and I data analysis is easy; we have spent thousands of hours with computers, SPSS, and analyzing data. It is not easy for someone who doesn't know the language (mathematics, specifically statistical terminology) and has no experience. The intense trauma experienced by these students was the major reason that we have invested so much time and effort to create a manual that is, above all else, clear.

This clarity is accomplished in several ways. First, we both come from the perspective that learning to use statistical software is best accomplished by *doing* rather than studying *about* it. That is why almost everything in the book is example-based. A few decades ago (60s I believe) the "new math" was introduced based on the assumption that if students understood *why* they were doing what they were doing, that it would enhance their ability to do arithmetic. Thus first and second graders were subjected to subsets, supersets, elements of a set, numbers in base 8, transitive property, reflexive property (remember that one? something is equal to itself) and a lot of material that rightly belonged in classes taken by first year math majors in college. The kids don't give a hoot about *why* 2+3 equals 3+2. It works, that's why. They can see it with sticks, and marbles, and dolls. The new math died a well deserved but relatively painless death soon after its inception, however, much of teaching in many classes today (particularly high school and college) is still based on the assumption that if you read enough about it you'll be able to do it.

I would challenge anyone who teaches research methods by reading and testing on some book for 15 weeks to match my students who read the book for 5 weeks and then spend 10 weeks taking a fairly major project from the initial idea to a finished APA-format paper. There is far less time reading about and much more time actually *doing*. The 15-weeks-of-reading students wouldn't have a clue of how to conduct a real research project. The knowledge comes from doing. This, of course, parallels the focus and intention of our book. It is a book designed to do. Some reading, of course, is necessary. However, after reading we urge that students practice, practice, practice on the computer, with real data, if possible.

With this perspective in mind the present book has been created. Every line of the book has as its focus to make clear both the conceptual idea behind the statistical procedure and the procedure itself. As mentioned before, the 10 data sets are available at **www.abacon.com/george** and may be accessed directly from that site or downloaded to a personal disk for future use. This allows students to practice the procedures they are reading about. How successful we have been in creating clarity is now out of our hands and remains for you to judge. We welcome your input and suggestions. Our addresses (both geographical and e-mail) and phone numbers are included with the book and your comments will gain a response.

Because Allyn & Bacon plans a new revision every year, your comments can be incorporated into new editions. Why yearly revisions? The present book is based on software, SPSS software to be specific. SPSS is like any other software business in the sense that they wish to maintain a competitive advantage. Because of this they upgrade their software regularly primarily to keep their programmers employed and secondarily to provide irritation to their users who have just become acquainted with the previous version. Often their changes are beneficial. For instance the shift from the command-style format of the PC+ version to the click-and-paste of the windows version was nothing short of brilliant. A bulky, terrifying, process suddenly became easy. Error messages, by far the most frequent output you would see with the PC+ version, essentially disappeared when you could click to select a variable rather than trying to remember how to spell a variable name.

Their revisions, however, also have a down side. The shift from the straight Courier text format version of output (Versions 6.x) to the present pivot tables in my opinion creates more confusion than benefit. With a small file and limited output, fine; but how often do you have a small file when you analyze data? Just for entertainment conduct a factor analysis and see if you can find anything in the bewildering pivot-table output that materializes. Our intent is not to be critical. We're sure that SPSS will tidy up some of these difficulties in new editions. . . . but I digress.

New revisions will happen. We plan to keep up with current editions of the software. If a new SPSS release doesn't come out in a particular year, a new edition of our book will be nothing more than correcting the occasional typo and incorporating your (or our) ideas to enhance clarity of the book. When a new version of SPSS software has occurred, then the revision will reflect the revised software. Allyn & Bacon plans to keep at least some of older editions for universities or individuals who have not yet upgraded to the most recent release.

## Types of classes for which this book is appropriate

The manual that you are now reading is designed to serve a much narrower audience than the textbook. While the manual addresses the needs of undergraduate statistics classes or research methods classes (across a wide array of disciplines), the book would also be appropriate for more advanced multivariate analysis classes, graduate statistics courses, major research projects (theses, dissertations), and as a reference guide to anyone who conducts data analysis occasionally or frequently.

I include fairly extensive analysis in the Chapter by Chapter Commentary section of this book. My comments about how to teach using this book must, therefore, be preceded by some description of classes with which it has been used. I have employed selected chapters from this book to assist in teaching my research methods courses. *SPSS for Windows Step by Step* (or prior editions of it) has, however, been my *primary* textbook in four different classes. One was a class on SPSS (Cal State Dominguez Hills) specifically devoted to the understanding of how to use SPSS software to analyze data. I have once taught an Introduction to Statistics course using the text (much to the dismay of the math department who thought it wasn't mathematical enough), and I have four times used the text to teach a course called Multivariate Analysis. Introduction to Statistics is a prerequisite for Multivariate Analysis and the class is largely devoted to actual analysis of data, or the practical application of statistical concepts in the real world. The latter class (and the actual SPSS class at the Cal State school) are the classes that provided the best fit for the book. Curiously, the Introduction to Statistics class was I think my greatest success teaching with this book. We managed to complete all the way through discriminant analysis with the lowest grade (in a class of 28) being C-. I felt this was testimony to the simple, direct treatment of material, and perhaps, to a lesser extent my fluency in the topic (at the practical level) and ability to present it clearly. Astonishingly, 10 students, after the class was over, actually purchased the SPSS software to install on their own computers! This suggests that they not only understood, but they actually wanted to continue using statistical software.

During my chapter-by-chapter presentation in the next section, when I refer to techniques of presenting material in my classes, I am typically referring to the Statistics or Multivariate Analysis classes. Those comments should be considered and implemented in your own classes with that in mind. If your goals in a particular class are quite different, many of the suggestions presented here may not be applicable.

## General Overview of the Book

An understanding of the structure of *SPSS for Windows Step by Step* should be useful. Chapters 1 through 5 are either introductory in nature or relate to SPSS processes that may apply to many of the statistical procedures that follow (Chapters 6-28). In the first five chapters special efforts have been made to present material in a way that is clear for the beginner. Of course, we have attempted to be clear in every chapter, but clarity in log-linear models, logistic regression, factor analysis and other advanced procedures is quite different from clarity in the first five introductory chapters or the 11 chapters of basic statistical procedures that follow. In the first 16 chapters we step-by-step carefully through the material and attempt to illustrate everything possible. The final 11 chapters deal with more complex procedures and a substantial grounding (viz., classes in advanced statistics) in those techniques is, in most instances, necessary. We still attempt, however, to make the conceptual bases of, say, factor analysis clear in the 6-page introduction, the many options straightforward in the Step by Step section, and the SPSS printouts coherent in the output section. We are under no illusion, however, that a beginner could understand factor analysis based on our presentation.

By contrast, we have made every effort that introductions to the first 16 chapters would give a clear conceptual feel for different procedures *even* for someone who had never taken a statistics class. As we move into two- and three-way ANOVAs and multiple regression analysis, we undoubtedly become less successful at our efforts, but the effort is always there. I feel that a student with good math aptitude (necessary!) even without a statistics course could make very productive use of the first 16 chapters of the book. We don't recommend it; we feel the book is most useful to persons who have already had statistics (or are taking it concurrently while learning SPSS), but this has not blunted our efforts to make it clear to the novice.

For the more advanced procedures, there was discussion with the publisher about writing a more limited book (for instance, one including only the first 16 chapters). We argued against this. The Introduction to Statistic class (mentioned earlier) progressed all the way through discriminant analysis. It was nice to have the same format description of the more advanced procedures to assist in the process. We know that most introductory classes will not go beyond chapter 16, but, if someone is working with questionnaire data and wants to conduct a reliability analysis, the chapter is there, painstakingly written to be as clear as possible. Same for MANOVAs, discriminant analysis, or log-linear models.

## A More Detailed Overview

The first section of this manual (pages 1-10) introduces the reader to the structure and content of the book. In the next major section of the book (Chapter by Chapter Commentary, pages 11-46) we will extend the brief presentations offered here to address the concerns of those who teach using the book. In those pages we go into much greater chapter-by-chapter detail. In the three pages that follow, we consider additional preliminary information concerning the structure and content of individual chapters.

**Chapter 1** is introductory. It is warm and friendly; it talks about SPSS, it describes the structure of chapters that follow. Most importantly, it introduces the example that is used in 16 of the 28 chapters and identifies typographical conventions that are used through out the book. When I teach a course using the book I require that students memorize the typographical conventions. It is not much of a burden because they are not numerous and are very intuitive, but it acknowledges the importance of knowing the language before you proceed.

**Chapter 2** deals with SPSS processes. While this chapter caters to someone who knows very little about computers and their use, there is much SPSS-specific information that would be useful to anyone. Included are topics concerning:

- mouse functions,
- taskbar and start menu,
- common buttons, both general and SPSS specific,
- the data window and other commonly used windows with icons identified and defined,
- the output window with its bewildering pivot tables, also with icons identified and defined, and
- instructions for printing output.

This chapter is often cited in the remainder of the book in case someone tries to use the book to conduct data analysis without reading Chapter 2. For instance, in the print sequence included in every analysis chapter, we refer the reader to Chapter 2 if they desire more information about print options or editing output.

**Chapter 3** concerns data entry and creation of data files. This chapter painstakingly step-by-steps students through conceptual ideas (how to design research and organize data) and procedures (how to name, format, and enter variables). It also considers editing already-entered data, and provides a complete listing of all data in the file (**grades.sav**) so frequently used throughout the book.

**Chapter 4** we regard as one of the most complex and important of the book. The ideas are not complex, the procedures often are. It would be rare for anyone sitting down to an analysis session to finish without making use of one or more procedures included in the fourth chapter. We deal with manipulation of data. Each concept is presented with conceptual rationale, examples to illustrate, and step-by-step instructions to clarify access. Included are the topics of:

- listing cases,
- replacing missing values,
- computing new variables,
- recoding variables (both switches of coding for existing variables and creating new variables via recoding),
- selecting subsets of data for further analysis,
- Reordering the data, and
- merging files.

**Chapter 5** deals with graphing procedures. The chapter is relatively brief (8 pages) and concerns itself with identification and editing of graphs more than with the multitude of graphs offered by SPSS. In many instances creating graphs is included within the context of particular chapters and students are referred to Chapter 5 if they wish to edit to enhance the appearance of the default graph. If students create graphs, however, much of the information presented is germane and will assist in editing and re-formatting of their graphs.

**Chapters 6 through 28** deal with different types of analyses. We need not list them here; you can read the table of contents. A word on organization, however, might be useful. SPSS is presented in different modules. The modules for version 10.0 have been dramatically revised since the 9.0 version: The Base System Module has been beefed up to include (in addition to those included in version 9.0) reliability analysis, multidimensional scaling, factor analysis, cluster analysis, and discriminant analysis. We cover ALL basic procedures presented in the Base System Module and several selected procedures from the Advanced Statistics Module and the Regression Models Module.

Our book has been organized around this module structure:

- The first 5 chapters are SPSS processes that apply to all modules,
- Chapters 6-22 apply to procedures offered in the Base System Module,
- Chapters 23-27 describe procedures from the Advanced Statistics Module or the Regression Models Module, and
- Chapter 28 (on residuals) applies to all three modules.

The module structure is not a trivial issue. Undoubtedly you are working with a network of computers at your college or university. Each module is purchased separately. Thus the issue of budgetary constraint comes into play. Often a school can afford only the Base system module. If so, then only Chapters 1-22 and Chapter 28 are applicable to the system you are using. The book is created so as to make it easy to manage if you have only one or two of the modules covered here.

A word on the structure of each of the analysis chapters (6-27) is important. Each of the chapters is self-contained. That is, armed only with an understanding of the typographical conventions, if a person wished to conduct t-tests on certain material, they could start with Chapter 11 and conduct t-tests. True, if they didn't know how to create a data file, step 1 tells them to go to Chapter 3 and learn how, if they get output and don't know how to edit the material they will be referred to Chapter 2; but otherwise, everything is there. Each of the Analysis chapters is divided into three sections:

1) **Introduction**: The introduction of each chapter explains the procedure as clearly and parsimoniously as we are able. For several of the introductory chapters the introductory material is only about one page long. We see no need to take more than two lines to explain what a "mean" is. Notice our brief section on measures of central tendency from chapter 7:

> The **Mean** is the average value of the distribution, or, the sum of all values divided by the number of values. The mean of the distribution [3 5 7 5 6 8 9] is
>
> **(3 + 5 + 7 + 5 + 6 + 8 + 9)/7 = 6.14.**

The **Median** is the middle value of the distribution. The median of the distribution [3 5 7 5 6 8 9], is 6, the middle value (when reordered from small to large, 3 5 5 **6** 7 8 9).

The **Mode** is the most frequently occurring value. The mode of the distribution [3 5 7 5 6 8 9] is 5, because the 5 occurs most frequently (twice, all other values occur only once).

Ours is not a statistics book or a commentary. We see no need to be anything other than clear. On the median we ignore the issue of "how about if you have an even number of values?" That's your job.

In other instances we go into some detail if we feel it is necessary for clarity. In Chapter 7 we spend a full page on statistical significance. More could be said about it, of course, but we felt it important enough to attend to it carefully. Also in chapter 7 we spend a full page on the normal distribution. All statistical procedures presented in the book (with the exception of Chi-square analyses and Nonparametric tests) are predicated upon an assumption of normality of your data. Thus, we felt it worthy of a full page.

On other procedures we may take several pages to inch our way through a clear presentation. Paul and I know that in our 28 chapters, some are better than others. I feel one of our best is the chapter on simple linear regression and curvilinear regression (Chapter 15). Regression is a difficult topic to understand and I feel the illustrations used to demonstrate regression in our book make a complex procedure essentially clear. That is the goal in all of our chapters.

2) **Step-by-Step sections**. Each Step by Step section begins with the identical first four steps (with minor variations depending on the procedure) and ends with the same two steps (explaining how to print their output and exit the program). It is the middle (various versions of Step 5) where the instruction takes place. There are as many step-5 versions as there are specific analyses presented in a particular chapter (labeled 5, 5a, 5b, 5c, etc.). Please refer to any of Chapters 6-27 to see examples of these. Prior to the step 5s, are descriptions of essentially all screens (or dialog boxes) used in that procedure. In this portion of the Step by Step section, a narrative description of the screens plus identification and definition of terms helps the student gain a thorough awareness of how to use the screens in the analyses.

3) **Output**: the output includes the results of analyses produced by SPSS. In almost all instances we use the format provided by the 10.0 version but edited to be clearer and more space conservative. This editing process has not been an easy one: The pivot tables as produced by SPSS are often awkward and cumbersome. In addition to printouts, we provide text commentary of what the output means. Finally, most output sections contain one or more boxes that define all output terms. These terms are also included in the comprehensive Glossary at the end of the book.

# Section II

# Chapter by Chapter Commentary

# Chapter 1: Introduction

## The First Four Sections, pages 1-5

The content of the chapter is clear from the Chapter cover page. The first four sections, (Necessary Skills For the User, Scope of Coverage, Overview of the Book, and Organization Chapter by Chapter) are largely for perspective and have a friendly, chatty feel to them despite the technical content. To the absolute novice, the Chapter by Chapter section would be essentially non-understandable and boring. The other parts, however, give one a feel for what is about to take place.

The final two sections (the example, and typographical conventions) is where we move away from friendly noises to material that is critical for understanding. We start with the example.

## Introduction to the Example, pages 5-6

As explained briefly in the introduction, the purpose of creating the example data file was to allow students to focus on data analysis rather than spending a lot of time trying to understand the file. In teaching a course using the book I actually quiz them on content of the data file to assure that they have reasonable fluency in the file that is used so often in the remainder of the book. For introductory students it is well to know that this file is used to illustrate procedures in all but two (Chapters 15 and 16) of the first 19 chapters. The data are all fictional, but do a good job of demonstrating many types of analyses.

## Typographical and Formatting Conventions, pages 6-8

This is the other section critical for understanding. We have created these conventions to be as intuitive as possible, and a thorough understanding of the 2 pages that comprise this section will place the users of the book in an excellent position to make most effective use of the chapters that follow. In my classes I will quiz students to assess their profound internalization of this material. The number of items is not great, but it sets a secure foundation for the remaining chapters.

# Chapter 2

Chapter 2 is divided into a number of sections, some more important than the others. Having used this book (or prior versions of it) several times to teach statistics-related classes, I will attempt to help you place in perspective the amount of attention each section deserves.

## The Mouse, page 10

This section is for the novice. It is worth reading (less than a page) for anyone who is not acquainted with mouse processes. There is, however, one element of substantive importance here, that is the right click: The right click often speeds up the data analysis process by its influence on variable names. A right click in many settings identifies the **variable label** and the **value labels** in a convenient box that disappears when you click elsewhere. How often have you tried to remember how you coded gender or ethnicity? This feature eliminates ever wondering. Also, a right click on many other terms will provide a definition.

## The Taskbar and Start Menu, Common buttons, pages 11-12

Both sections are essentially for the novice, but serve as a handy reminder for anyone who has forgotten.

## The Data Window, page 13

This is of substantial importance in the early stages. Once the data window has been used a few times, its importance diminishes. The lists of icons are handy but hardly worth memorizing, besides, the page-14 illustration is one of the front cover screens, so if a student wishes to review the meaning of the icons, they are always handy. True, when you place the cursor on an icon it identifies its function. The advantage of the descriptions on page 14 (and the front cover) is that they often have more complete definitions. Most frequently used icons include:

- Open file
- Save current file
- Print file
- Undo the last operation (this has no memory, it undoes only the most recent operation)
- Go to a particular case number
- Access information about the current variable
- Find data
- Insert subject or case into the data file
- Insert new variable into the data file
- Shift between numbers and labels for variables with several levels

## The Open Data File Dialog Window, page 15

This is the second of the front cover screens. It is intuitive and not very difficult, but this is one spot that it would do well for the professor to stop and spend a few minutes. The four methods for accessing a data file listed on page 16 are important for users to internalize.

## The Statistical-Procedure Dialog Windows, page 17

Every statistical procedure has a main dialog window. Many of the features are similar in each window. The OK, Paste, Reset, Cancel and Help are always included. Also, the list of available variables is always present. This section is worth some quiz items because of its centrality to all statistical processes.

## Keyboard Processing, Check Boxes and Radio Buttons, page 18

The check boxes and radio buttons should occupy a solid 15 seconds of your lecture. The keyboard processing is not as trivial as it might initially seem. Few people are scared of mice, however, many times the key strokes are far quicker than the equivalent mouse processes. For instance when creating value labels, once you have identified the value and its label, you may click "Add" or press "ALT-A". The ALT-A is much faster. Talk with students about the meaning of the underlined letter.

## The SPSS Output Navigator, pages 19-25

This topic occupies 6 pages because (in my opinion only, I'm sure) the pivot tables SPSS has introduced are such a nuisance. With small data sets or very limited output they are kind of cute. The pivot features allow you to switch rows and columns, put labels on their ends and to engage in a variety of other entertaining activities. With large outputs, they become a special type of purgatory. In several later chapters (more complex procedures) SPSS drops pivot tables entirely and goes with their linear, straight text format. I personally find the text output much easier to deal with.

How you respond to this section of the chapter is, of course, your prerogative, but it cannot be ignored. A good starting point is to read the 6 pages yourself and practice with some actual output. You can click and drag the lines of the output to create a neater look. SPSS has not yet figured out how to get value labels to fit in the boxes, but some click and drag activity can create greater order. I haven't quite decided if my irritation with present format is due to the thousands of hours I have spent with the simplicity of the former versions' text output, or if the present format is truly more cumbersome. Here, however, are some tips about issues that need to be covered:

Page 20 and 21: Distinguish between the two halves of the screen: the Outline view (to the left) and the actual output (to the right). The heavy border between them may be moved (via a click and drag mouse operation) so as to eliminate either section and allow a clear version of other window. Still on page 20, frequently used icons include:

- Open file
- Save file
- Print output
- Undo the last operation (it has no memory, it will only undo the most recent operation)
- Go to the SPSS Data editor (very handy to access your data screen on demand)
- Go to a particular case number
- Get information about variables
- Display currently selected object (a double click on the closed book icon in the outline view to the left will accomplish the same, see diagram on page 21)
- Hide currently selected object (a double click on the open book outline icon in the outline view to the left will accomplish the same, see diagram on page 21)

The small + or – signs to the left of outline items (see visual display on the top of page 22) are also important to be aware of.

The **pivot functions** should at least be presented: Explain that it first takes a double click on the output object to activate the pivot tables. A new menu bar heading will emerge called "**Pivot**" that allows you to access a number of pivoting functions. Beyond that I would suggest that you tell your students to play around with it to see what each menu item does. Warn them to save their file prior to their experimentation so if they tie themselves in hopeless knots, they can simply revert to the saved version of the file; or, simply run the analysis again. The "persistence" function of SPSS main dialog windows allows an analysis to be rerun without having to remember which variables or options you included on the previous try. If you have a burden to help your students become fluent in pivot-table operations, you will need to create your own agenda.

## Printing Output, page 24

Printing output is quite straight forward. The print dialog boxes are intuitive and very similar to many other print dialog boxes for word processing or other software programs. The only issue of significance concerning printing is the frequent need to edit the output prior to printing. SPSS has a long history of collusion with the timber industry to consume as many pages as possible in their output. Thoughtful editing prior to printing output can save many reams of paper. In my classes I will often specify: "Edit output so it fits on one page." This gives students practice in output editing, and, by time the course is finished, they will have become fairly fluent in this practice through many instances of actually editing output.

Printing a *data* file is quite a different dynamic. There is really no space-conservative way that either Paul or I have found to print an entire data file. SPSS is quite capable of printing even a large file but it takes many pages. The key concern is that usually a researcher wants to print only a *portion* of the file. This is accomplished by highlighting the desired material: Click & drag within the data file to print a particular rectangle of data; click on the variable name(s) to print particular variables (all cases); click on the case number(s) to print certain cases (all variables). When you arrive at the print screen, make sure that the "Selection" option is highlighted prior to clicking the OK.

# Chapter 3: Creating and Editing a Data File

While the previous chapter provides material that may or may not be important to cover in a class setting (depending on the expertise of your students), the data-file chapter requires a thorough treatment for just about anyone. If students become aware of the importance of carefully-crafted research design prior to entry of data, coherent order and formatting of variables in the data file, and certain modifications and manipulations that may be enacted with these data; they will have a strong foundation to negotiate the remainder of the class successfully. In my experience teaching classes using this book, this chapter cannot be either overlooked or minimized. I typically quiz thoroughly on all items I feel are foundational to creating the data file. In addition, we assign a good deal of practice on the computers that consists of actually creating a file and entering data. Now, it is true that any intellectually-challenged human can enter data with minimal instructions. To construct a meaningful data file requires intelligence, experience, and finesse.

A typical assignment directs students to recreate the data file for the **grades.sav** file (described in Chapter 1 and listed at the end of Chapter 3). Once students have created, named, and formatted all variables, then they enter data for the first 10 or 20 subjects from the lists at the end of Chapter 3. After completing this assignment, students (in a hands-on intensive course) will have many opportunities to edit variables (change coding, cut and paste, create new variables) and perform other manipulations with the data file. I have found that a thorough understanding of the contents of Chapter 3, exercises such as the one described above, and a course that gives students lots of practice working with data files, provides a secure foundation for creating their own files in the future. Some specific areas of interest follow.

## Research Design and Structure of the File, page 28

Instruction concerning the nature of a well-designed study is central to conducting a research methods class, regardless of the particular discipline (e.g., psychology, sociology, political science, etc.). In the text we underline the necessity of a carefully designed study but are in no position to go into much detail. We would encourage you to expand and elaborate on the material introduced in the book. If, on the other hand, you are teaching a statistics class, the best that can be done is to indoctrinate students in the reality that carefully crafted research is central to effective data analysis. The actual procedures of how to do this will probably wait until a research methods course.

## Step by Step, page 29

The screen shown here is a slightly modified version of Front Cover Screen 1. It is important to know, of course, but is so intuitive that fairly limited experience will provide all the introduction students need. If they ever require visual reference during the course of data analysis, they can always check the front cover, which also includes the meaning of each of the icons.

## The Variable View Window (Screen 3.2), page 30

This screen represents a dramatic revision from all prior versions of SPSS. We applaud SPSS for these changes. A method of formatting data, which we felt was already good, is now even better. It allows equal fluidity at first typing all the variable names and then formatting each one afterwards (not available with Version 9.0), or, typing a single variable name and selecting all formatting options before continuing to the next variable. When a variable name is typed and no formatting information is provided, SPSS will automatically assign that variable the default settings (numeric, 8 characters wide, 2 decimals, no labels or values, etc.). If this is acceptable, fine, if not, changes are as simple as clicking on the desired cell and making those changes (described in detail in the text book). Finally, if the individual desires to make sure that values enter correctly, he need only click on the Data View tab, enter a value (under the variable name now listed at the top of the screen) and see that the value is entered and formatted correctly.

## Step-by-Step Sequences. pages 31-35

These seven step-by-step boxes identify in lucid format exactly (click at a time) how to format several different types of string and numeric variables. I have found that these five pages are sufficient for students to learn the basics and I rarely have questions from students about variable formatting after that. Once again, SPSS has made this quite intuitive so a little bit of learning will go a long way.

### Entering Data, page 36

This section begins with how to save (initially) and save (subsequently) the data file. You will avoid the occasional panic attack or epileptic seizure if you remind students to frequently save data during the entry process. There are several ways to enter data. A favorite method of Paul and myself is, following the creation and formatting of EACH variable, enter the actual value for the first subject—just to make sure it fits and is formatted correctly—then go on to the next variable. After the file has been created and formatted (and one value has been entered for each variable), then go on to enter the remainder of your data. Whether this is done row by row (cases/subjects) or column by column (variables) depends entirely on the way your data are formatted.

## Editing Data, page 36

Changing the cell value, inserting a new case, and inserting a new variable will be utilized thousands of times by any serious researcher. To change an entry is as simple as clicking on the desired cell, typing the correct entry, then pressing the tab, enter, or one of the cursor keys. Inserting new cases or new variables requires a little more attention. A click on either of these icons clears a column (or row) before the selected column (or row). You may then either type in new data (with a new variable name when appropriate), or you may paste a new row or column into the recently-cleared space. It would be well for you to pay special attention to the material on page 37 to ensure that your students cut and paste cases or variables correctly. Searching for data is intuitive, straight forward, and very handy. This is particularly useful for replacing errors in your data file.

The chapter concludes with the data included in the **grades.sav** file. Be aware that the file is available at **www.abacon.com/george** and may be accessed directly from that site or downloaded to a personal disk for future use. includes four additional variables (**total, percent, grade**, and **passfail**). All information necessary to create the **grades.sav** file is included on the last three pages.

# Chapter 4: Managing Data

As suggested by the first paragraph of this chapter, we consider this as important to the mastery of statistical analysis as the creation of a data file. Let us be frank: This is also one of the most difficult chapters for students to master. The ideas are not at all difficult conceptually, but they are often troublesome to implement *and* too hard to remember a month or two later.

## Case Summaries, page 43

In former versions this was a simple "List Cases" function that allowed you to select certain data (any number of cases/subjects and any number of selected variables) to print out in a nicely space conservative (72 lines per page, 10 point Courier font) format. Now SPSS has expanded the command (as reflected in the name change to "Case Summaries") to include a variety of other options (not really a problem, you don't <u>have</u> to select them), but have reverted to the awkward boxed output (allowing no more than 30 lines per page). I don't express this mild cynicism in the book, and the command it still useful. I would guess that most people who use SPSS are not even aware of this option, but it proves to be convenient in a variety of settings, particularly to ensure that data are entered, formatted, ordered in the desired manner. I give students several problems involving case summaries to heighten their awareness of its importance.

## Replacing Missing Values, page 46

A read through this section will reveal the bias that both Paul and I possess. We feel that missing values should be dealt with at the data entry phase, particularly for categorical variables. If we do replace missing values with the mean value for the distribution, we are more likely to base the replacement on the mean of other values in the same domain FOR THAT SUBJECT rather than for the average for all subjects in the data set. For serious research leading to publication, I will consistently use Method 3: Create regression equations to produce predicted values and then enter those values into the original data set. This process is certainly beyond the resources of most introductory students; more advanced students, however, should not have difficulty with it. Having identified my bias, let's continue.

Distinguish between the terms **pairwise** and **listwise**. Students will see these terms frequently. Also clarify the distinction between "system-missing values" and "user-missing values", something beginners are not likely to use, but the terms appear from time to time. The 15% number is important to incorporate into their thinking. The last paragraph on page 48 is crucial. Stress to your students that it is not possible to solve the problem of missing values by simply clicking boxes. What we demonstrate in sequence step 5b provides something that is statistically viable but is only obliquely related to what the student's actual GPA or quiz score might be. I do not typically provide exercises in the missing value section. I do, however, cover the material and quiz on it.

## The Compute Procedure, page 48

This material will be frequently used and should be carefully taught. Some hints: Instruct the students that creating a new variable will produce a value (based on their formula) for every subject in the data set. This new variable will be given the name they designate, and will appear, after computation, at the end of the data file. This may then be cut and pasted into a more convenient location in the file if they wish.

- Teach students to double click on the variable name (in the list to the left) to paste it into the active box rather click the name and click the right button. It is quicker.
- For introductory classes, avoid the 70 functions all together. They are so idiosyncratic that a perusal of the 70 found that I understood only 16 of them (and I have an undergraduate degree in mathematics!).
- I would suggest that you have students compute 7 or 8 new variables to learn the process thoroughly. The complexity of the formulas to create these variables depends on how advanced your class is. I would urge that you provide simple examples (such as sums of quizzes and final for total points, or computing percents) so they become acquainted with the process before they are forced to consider complex computations.
- The random number function (bottom of page 50) can be quite handy, however, it extends beyond the resources of an introductory class.
- The step-by-step examples on page 51 are excellent for first computations. Just remember that if you are creating these off the **grades.sav** (available at the **www.abacon.com/george** web site), you will need to give the variables different names than the names used in the original file (e.g., "total1" rather than "total"). Once completed, they can check their results with equivalent values in the data file to see if they did it correctly.

## The Recode into Different Variables Procedure, page 52

An important skill, but not one worth memorizing. I have my students work the examples on pages 53 and 54. Both give an excellent feel for the most frequent uses of this function. The command is not particularly intuitive and may require some work to master. My experience is that working the two examples is enough. If students, at some future time, need to compute into different variables, then they will have to go back and review pages 52 and 53 to remember how. Although this command is not used all that frequently, it is well to at least remember where in the book they can access the step-by-step information.

## Recoding into the Same Variable, page 54

This operation is used much more frequently than its brother in the previous paragraph. It is also much more intuitive in terms of the step-by-step procedures. Several exercises of changing the coding are quite useful; sequence step 5g on page 56 provides an excellent starting point. The final paragraph on page 56 is not to be ignored. It is mandatory that if you change the coding that you will also need to go into the variables dialog box and change the value labels associated with each value. As noted in the book, some horrifying misinformation can result if the labels are not changed as well.

## Select cases, page 56

This is, hands down, the most frequently used procedure in the chapter. At a general level it is important to teach the process: Provide them with several problems that give them practice, and be sure to include key concepts from this section on quizzes. However, once learned, they will use this option so often that before long it will become second nature.

The most critical concern is determining how to select complex combinations of cases. For instance, in a data file on spirituality (not included on the disk), there are 13 different religions listed and coded. If we wished to do a Catholic versus Protestant comparison this might be difficult because there are 7 different Protestant groups but only one Catholic group. I would suggest that you present introductory students with only fairly simple select cases options. If you have a more advanced class deal with more complex selections when you encounter them in other analyses. It is necessary that you proceed with caution when using the logical operations (bottom of page 49). The ampersand (&) and the pipe (|) can easily be mixed up. Best bet is for students to create their statements, then check the data file to see if cases have been selected (or deselected) correctly. If not, go back and try again. In time this will become intuitive.

Finally, the second to the last line on page 57 (selection of sophomores and juniors) provides an example of an idiosyncrasy that students must comprehend if they wish to avoid major frustration. The intuitive first impulse would be to code for sophomores and juniors by a "**year** >= 2 & <= 3". It won't work. You must state the variable name (**year** in this case) after each logical symbol: "**year** >= 2 & **year** <= 3".

## Sort cases, page 58

This is a mercifully simple little procedure, and, very very practical. Frequently you might wish to view your data set ordered from high to low (or low to high) on some particular variable. This is often used for class lists: alphabetical by class, alphabetical by section, from high to low based on total points or percents, from low to high based on student ID numbers. Not much effort for either you or the students.

### Merging files, page 59

This process is beyond the scope of an introductory course. Indeed, for complex files, it is beyond the scope of just about anyone. For advanced classes, there are two files (**graderow.sav** and **gradecol.sav**) that are formatted to neatly illustrate this procedure.

# Chapter 5: The Graphs Procedure

This is, to some extent, an in-transition chapter. In earlier editions of this book, the chapter was extensive and covered many different types of graphs. The shift (in this version) to a leaner, overview-and-editing-options type of chapter was for essentially two reasons:

- In the earlier edition, some of the graphs we presented were context-specific and thus lost clarity if the reader was not acquainted with the procedure on which the graph was based. For instance, we showed how to use multiple line graphs and clustered bar charts to represent ANOVA interactions. If a person had never heard of ANOVA, it was a somewhat pointless exercise.
- The other difficulty is that there are so many different types of graphs and different options for each graph that it would make the present chapter bulky and cumbersome to cover even a fraction of graphs available. The SPSS Manuals spend 165 fine-print pages on charts alone. For the present, we would suggest that you consult the manual if you wish to present a specific chart of interest. If you, as an instructor, wish to present particular types of charts in detail, you would need to create your own agenda from material presented in the manuals. On the other hand, a number of frequently used graphs are presented in our book within the context of the Chapters that make use of such graphs (e.g., bar charts and histograms in Chapter 6, line graphs in Chapter 14, scatter plots in Chapter 15 . . .). We anticipate expansion of Chapter 5 in the 4th edition to cover some of the graphs not presented within specific chapters. Any suggestions from you (via e-mail—addresses in the front of the text book) would be considered before revision.

We have included instructions about how to access several different types of graphs in appropriate chapters; these include:

- Bar charts, histograms, and (somewhat obliquely) pie charts in Chapter 6.
- Multiple-line graphs and clustered bar charts in Chapter 14.
- Scatter plots showing both linear and curvilinear trends in Chapter 15.
- Predicted-probability charts in Chapter 20.
- Scree Plots in Chapter 24.
- Icicle posts and dendograms in Chapter 25.
- A number of charts involving residuals in Chapter 27.

The primary function of Chapter 5 as it is currently constructed is to provide:

- An overview of the graphics capabilities of the current SPSS graphics editor.
- A sample graph that identifies almost all of the SPSS-specific graphics terms.

- Two types of dialog boxes that occur in just about every graphics procedure offered by SPSS when accessed through the Graph command.
- Two sample graphs to demonstrate graphic output.
- An identification of how to edit the graph (starting with a double click on the graph).
- Identification and description of each of the graph edit icons.
- Explanation of how to access and use other edit options.
- Description of key commands and options.

We feel the chapter does a good job (within 8 pages) of clarifying a number of important concerns. Prior to your students creating graphs, have them read through the chapter. How much time you wish to spend on this topic depends, of course, on the nature and goals of your class. I have covered the spectrum in classes I have taught from an extensive unit on creating and editing many graphs (extending over several class periods) to entirely ignoring the topic.

# Chapter 6: Frequencies

This is the first chapter that actually begins to look like data analysis. You and I know that it isn't, but some eager students begin to get excited as they start to do things that actually produce output. Within that context, we are now into the standard chapter format:

- The Introduction (pages 74-75).
- The standard 4-step sequence that accesses the program of interest (page 75).
- The step-5 sequences (pages 76-80) that actually perform different types of analyses, six of them in this case (5, 5a, 5b, 5c, 5d, and 5e).
- The printing and exiting program sequences (steps 6 and 7) that complete each Step by Step section of Chapters 6-27 (page 80).
- The Output (pages 81-82) composed usually of SPSS style output, definition of terms in neat boxes, text that gives a narrative description of what the output means, and sometimes charts or graphs to help clarify.

Frequencies, bar charts, histograms, and percentiles are described with an absolute minimum of adornment. If you feel our brief presentation is inadequate, feel free to expand upon it. In the Step by Step section, once students have executed the first three of the introductory steps a few times, they will be internalized and Step 4 will be the initial point of interest. There is not a lot worthy of comment in this chapter because it is so straightforward. The initial dialog box (Screen 6.1, page 76) is simple and intuitive. Examples and exercises are available in just about any data set. Students learn this quickly.

A point that often causes frustration early in the process, however, is the order of variables in the initial dialog window. Variables in the box to the left may be ordered according to the "variable list" (the order they were initially entered), or "alphabetic". The variable list is typically the preferred order unless you have a very large number of variables, and all material presented in the text assumes the variables are listed in the order of the original variable list. If your students' variables are alphabetic and they wish to change, have them click **Edit**, then **Preferences**. There, in a box in the upper right hand corner, the two selections (variable list and alphabetic) are offered. Once the desired order is selected, all further dialog boxes will list variables in the designated order.

In this chapter (and also Chapter 7) we actually recreate the SPSS Output Navigator (the output screen) within the chapter for visual reference (with the output from sequence step 5 displayed). After Chapter 7 we simply refer students to Screen 1 on the inside back cover, a sample output screen that includes definitions of all the icons.

# Chapter 7: Descriptive Statistics

The first paragraph identifies all the terms that will be covered in this chapter. The most important ones from a statistical viewpoint are:

- Mean, median, mode
- Variance and standard deviation
- Skewness
- Kurtosis
- Maximum, minimum, and range

The standard errors are described in the narrative but are not considered central for an introductory class. Also the median, mode and size of the distribution are more frequently used in business and economics than in the social sciences.

New to this edition is a full page on statistical significance. In addition to the presentation here, significance is often addressed (more briefly) in the chapters that follow. Obviously this is an important concept in data analysis and you may wish to build on the material we present on pages 84 and 85, or you may wish to present this topic with the aid of another book or based on your own resources. The page is not designed to be comprehensive, but, to the instructor, it is a reminder to spend serious time on this issue.

The next topic, more germane to the subject of the present chapter, is a full-page description of the normal distribution. We attempt to make the concept clear by many illustrations. What I notice now (with mild horror) is that we apparently wait until the chapter on nonparametric procedures to indicate that the reason the normal distribution is so important is that most statistical procedures are based on the assumption of normality of data. Just as you don't build a house with rotten lumber, you don't build an analysis on variables that are not psychometrically sound (viz., normally distributed). We make strong allusion to this necessity in the blurbs on kurtosis and skewness, but in the next edition, the point will be drawn more directly.

Once again, the description of the terms and procedures are made with a minimum of adornment. You are free to expand if you feel such is necessary. The simple 2-line descriptions and simple illustrations seemed adequate to us. Kurtosis and skewness both merit an entire paragraph and a couple of graphs each. These two constructs provide the scales that actually measure the normality of any distribution.

For assignments in this material, I will typically have students print out psychometric properties of 40 or 50 variables in some large data set. I include only mean, standard deviation, kurtosis, skewness, and N so that the information about each variable will fit on a single line. They will then circle, underline, or box variables that are excellent for most psychometric purposes (kurtosis AND skewness between ± 1.0), those that are generally acceptable for further analyses (kurtosis or skewness between ± 2.0 but at least one of them is outside the ± 1.0 range), or those not acceptable (either kurtosis or skewness outside the ±2.0 range). This is a bit more mechanical than how a real researcher might use these measures, but it conveys the point clearly.

In the Step by Step section, the two dialog boxes (pages 89 and 90) are simple and intuitive. Students I have taught have never had difficulty with either of them. The output is straightforward, and the interpretation is clear once students have internalized the meaning of terms from the introductory portion of the chapter.

# Chapter 8: The Crosstabs Procedure and Chi-Square Tests of Independence

Ah! Now we move into material that is capable of actually confusing students! It is also the first genuine analysis chapter. I have found that as essentially simple as chi-square analyses are, it takes awhile before students are able to comfortably internalize some of the concepts. The widely used-used terms “independence’ and “dependence” I think a little unfortunate. They do not quickly convey what a chi-square analysis is attempting to accomplish. Paul and I have gone through tortured efforts to make the basic concepts clear in the introductory two pages. Based on my teaching experience (using the book) we have been only moderately successful. What follows are some thoughts dealing with how to most effectively convey this material.

First, you have taught these concepts before. Expand, modify, or replace what we have presented here by that which has worked best for you. If our description proves useful, build on those ideas.

Deal with the categorical/continuous issue. While you may perform crosstabulation with either continuous or categorical data, it is rarely acceptable to perform chi-square analyses on categorized continuous data for the reasons delineated in the first paragraph on page 94. Emphasize that with chi-squares you are always dealing with frequencies (or expected frequencies) of *categorical* data.

Next explain that what is being compared in chi-square analyses is whether **actual** (or observed) **values** differ significantly from **expected values**. Some more time may need to be spent with “what are expected values”. Whether you show the method for actually computing expected values (row percent × column percent × number of subjects) is up to you. We explain what expected values are in a seat-of-the-pants fashion in the introduction. We do not show the actual computations.

To demonstrate how the chi-square procedure really works, I have students actually compute chi-square values based on the formula shown on page 95. This allows them to see that many small discrepancies result in a small (and non significant) chi-square value, and that several large discrepancies inflate the chi-square value resulting in significant differences between observed and expected values. It’s not a difficult formula and gives students a good understanding of the procedure.

Concerning the explanation of dependence and independence: Try what Chris Cozby suggested in his research methods book: Use the “that depends” phrase. Are there more men or women in different classes? That depends on whether the class deals with, for instance, engineering (typically more men) or psychology (typically more women).

Are there more Christians or Jews in different classes? That depends on whether the classes are conducted in the US or in Israel. These four instances represent examples of dependence. The membership of the class (whether gender or religion) depends on certain characteristics of the class or setting. Independence might be illustrated by some obvious examples: Is gender balance dependent on ethnicity? No, we know that there are roughly 50% males and 50% females in any ethnic group. Thus gender and ethnicity are independent of each other. If this is helpful, good; if not, use your own examples.

To further illustrate, run students through several chi-square analyses and then interpret what they mean. On page 66 in this manual the **helping3.sav** file is used to demonstrate that there are significant differences between men and women concerning the situations in which they are more likely to provide help for a friend. Men report more instances of helping with goal-disruptive problems whereas women report more instances of helping with relational problems. This is a significant and interesting gender difference indicating that women are more likely to help in situations where emotional support is facilitative and men are more likely to help in situations where they can *do* something to fix the problem.

It may also be important to briefly address the correlation statistic that is typically printed out when conducting chi-square analyses. Here it is necessary to consider the nominal data versus ordinal data issue. Our book address this concern only briefly (in Chapter 3), but it is worth explaining that if there is an intrinsic order to the levels of your variable (e.g., levels of income, years of education) then meaningful correlations may be computed. If your variables are nominal (or even one of them is nominal) then correlations are meaningless. There is no intrinsic order to levels of ethnicity or different types of religions. The difficulty, of course, is that if you are teaching an introductory class, they have not yet been introduced to correlations.

The material on Phi and Cramer's V will typically extend beyond the level of an introductory class. However, if you wish to compare the magnitude of various analyses, then both are important. There are additional chi-square-related issues not covered here, but the chapter does an adequate job of addressing many of the most important concerns. There are a number of chi square analyses in the exercises section that you may use to illustrate this procedure to your students or to assign for homework.

# Chapter 9: The Means Procedure: Subpopulation Differences

This is a simple little procedure that is not really a form of data analysis, it is just a listing in crosstabulated format of certain characteristics of your data. While the Crosstabs procedure from Chapter 8 indicates the frequency (or the number) of subjects in each category or sub-category, the Means procedure, instead of listing the frequency might list the mean values within each category, or the standard deviations, or the frequencies if you like. For instance, you might want to crosstabulate a **section** by **year-in-school** on **total** points for your classes. This would produce the mean number of total points for students in each class (3 levels), for each year in school (4 levels). It is mainly a device for organizing data. The Means procedure does include a simple one-way ANOVA in their options, but it seems foolish to consider that issue here when Chapter 12 does a fairly thorough and thoughtful job of covering one-way ANOVAs.

How often is this process used? I never use it. I find that the Descriptives output does as well. Paul says he uses it a lot. I suppose it is a matter of preference. In my classes I present the material in five or ten minutes, have two or three simple exercises to conduct with it, and then forget it. I am sure that the emphasis placed on this chapter will be entirely dependent on the preference of the instructor. There have been documented cases of individuals who have actually lived full, happy lives, and have never even *heard* of the Means procedure.

# Chapter 10: Bivariate Correlations

If Chapter 9 can be lived without, Chapter 10 most assuredly cannot. Correlations are a central form of analysis in the majority of research projects, and, a number of advanced procedures are based on correlations, such as simple and multiple regression analyses, logistic regression, discriminant analysis, and structural equation modeling.

Paul and I frankly acknowledge that some chapters are better written than others (we will never acknowledge, however, that a particular chapter is worse than others), and we feel that the correlation chapter does a good job of presenting the material in a concise but lucid format. On pages 112 and 113 the explanations for the perfect positive, positive (but not perfect), no correlation, negative (but not perfect), perfect negative correlations with their associated graphs make clear what correlation is and how it relates to the associated phenomena.

In addition, we cover several other issues that are critically important to an understanding of the use of correlations. The brief paragraph or two that we present on each of these issues (pages 114 and 115) is largely a reminder to instructors that this material needs to be considered and we leave it to you to determine how. Although we introduce linear versus curvilinear issues here, we wait until Chapter 15 for a more thorough discussion. The topic of significance is once again presented briefly. Four compact paragraphs present the central arguments concerning the issue of correlation and causation, and partial correlation is considered briefly, but is covered in more detail in the ANOVA and regression chapters.

Additional issues that are covered in the Step by Step section include a Pearson versus Spearman contrast. Generally Pearson correlations should be used with continuous data and Spearman with categorical (ordinal, of course) data. The simple reality, however, is that the Pearson formula works quite well with ordinal data, and when there is a correlation between a continuous variable and an ordinal variable, Pearson is typically used.

Another concern is the selection of one-tailed or two-tailed tests. The rule of thumb that governs a choice of one-tailed versus two-tailed is that if you have a clear idea as to the **direction** of the relationship (a positive correlation or a negative correlation) then it is acceptable to use the more powerful one-tailed test. If you have a matrix of correlations in which you have little idea of the direction of relationships, then 2-tailed tests are more appropriate.

A concern addressed in the regression chapters but not included here is the issue of linear dependence. Even though our book doesn't consider linear dependence in the correlation chapter, I address the topic in my classes and it is a prominent issue in the exercises that my students complete. Often as an assignment I will have students compute a correlation matrix of a number of variables. Then I will have them identify correlations that are meaningless (I purposely include some categorical variables that are nominal), identify correlations in which there is numeric linear dependency (like quiz1 with total points, or final with total points—total points is the sum of the quizzes and the final), and identify correlations that may have conceptual linear dependency (such as anxiety and tension, outgoing and extroverted). To print out the matrix is no trouble for the students, but to identify the nature of the correlations often produces some productive anxiety and some constructive thinking.

Another element of assignments is for students to identify the five highest valid correlations (positive) and the five strongest (negative) correlations and describe what they mean in appropriate language (that is, non-causal if causality cannot be assumed). These might include sentences such as:

- Correlations showed that greater self-efficacy is significantly associated with more time spent helping ($\underline{r}$ = .37, $\underline{p}$ < .001).
- Results demonstrated a significant positive correlation between spirituality and life satisfaction ($\underline{r}$ = .17, $\underline{p}$ = .023).

We introduce the frightening foray into SPSS syntax to create a correlation matrix that is not symmetric. The inability of SPSS to compute asymmetric matrices (without using syntax) is a distressing failing of all the windows versions. To demonstrate its utility: If you have computed a 10x10 matrix of correlations and then add two new variables, you may wish to compute correlations of those two with the other 10. The material presented on page 118, I feel, gives a clear picture of how to do this. Those who have used the PC+ version of SPSS have encountered these syntax command files many times.

# Chapter 11: Independent Samples, Paired Samples, and One-Sample t Tests

T tests represents another mainstay of data analysis. We regard this as another well-written chapter and my students who actually read it rarely have difficulty with *t* tests. The differentiation between independent-samples, paired-samples, and one sample *t* tests is presented simply and logically. I have found that additional words are rarely necessary in my classes to clarify further.

A topic suggested by the first sentence of the chapter recommends a discussion that might take place at this time if it has not already been presented. This is the distinction between populations and samples. In an earlier version of this book we had written the first sentence wrong and a reviewer (who was a statistician) tidied us up quickly. The issue is NOT whether two sample groups differ significantly, the issue is, do the *two populations* from which the samples were drawn differ significantly. This is central to the idea of statistical inference. You could venture into the identification of populations (Greek letters), population values (the unknown and unknowable real means, standard deviations, etc.) and indicate that samples (English letters) attempt to determine or estimate actual population statistics. I have found this juncture a good spot to discuss these topics.

I also find this a useful point to illustrate the dynamic of within-group variation as compared to between-group variation. Draw some normal distributions on the board and show how manipulating the distance between means or an increase or decrease of variance will influence the likelihood of statistical significance between groups. You might even (as an introduction to Levene, described below) indicate how unequal variances affects statistical significance. I usually present this material at a conceptual level, but when we get to one-way ANOVAs I have students actually calculate sums of squares to get a real feel of what the statistical procedure is accomplishing.

The introduction also briefly discusses tests of significance and one- and two-tailed tests. My students generally are fairly conversant on both these issues by this point in the course.

The content of the Step by Step section I have found to be straightforward. Students have acquired this material easily, and I have no particular comment about it. The output, however, presents some concerns. One issue is the use of Levene's test for equality of variances. It is not a very hard concept: If the variances differ significantly ($p < .05$) then use statistics based on the unequal-variance estimates. If the variances do NOT differ significantly ($p > .05$), then use statistics based on the slightly more powerful equal-variance estimates. It sounds (and is) simple, but slower students still mix it up.

In assignments with t tests, I typically have students print out many of them (for instance gender differences on all appropriate variables in a data set). Then I ask students to delete any analyses that do not show statistical significance, print out the ones that do show significant differences, circle the correct *t* value (based on the results of the Levene test) and then write up the results. The write-ups are simple sentences with the APA-correct statistical line: For instance:

- Men ($\underline{M}$ = 103.45) scored significantly higher than women ($\underline{M}$ = 98.32) on the final exam, $\underline{t}$(103) = 3.47, $\underline{p}$ = .03.
- The helpers rated the quality of help significantly higher ($\underline{M}$ = 5.42) than did the help recipients ($\underline{M}$ = 4.96), $\underline{t}$(535) = 4.56, $\underline{p}$ < .001.

I have found that students do fairly well at writing the statistical part of the output but have great difficulty writing a coherent English sentence. This makes grading the assignments quite a frustrating chore. Don't worry, though, it will get far worse in the ANOVA chapters.

An additional feature I include on home works is to have students highlight the dependent variable of interest, circle the two mean values, and circle the correct significance value. On paired-samples t-tests, correlations between the two levels of the variable are included. Since the topic of correlations has just been covered (previous chapter), students now understand the procedure and I occasionally have them write up the meaning of the correlations as well.

In general I have found t tests to be a straight forward and satisfying topic to present. The transition to ANOVAs, however, is another story.

# Chapter 12: the One-Way ANOVA Procedure

This chapter begins its polemic with a comparison of *t*-tests (exactly two levels) with one-way ANOVA (two or more levels of the independent variable). We clarify that one-way ANOVA includes exactly one continuous dependent variable and exactly one categorical independent variable. Here, then, is where description of ANOVAs becomes difficult. It is fine to say "if you conduct an ANOVA all pairwise comparisons are considered and a test statistic results that identifies if there are ANY significant differences within pairwise comparisons. If there are, then you do a post hoc test to identify where significant differences occur." It is a fine, noble sentence and largely correct, however, you get caught with egg on your face when a student computes a one-way ANOVA that shows a *p* value of .11 and yet ends up with significant differences in pairwise comparisons. The same thing happens with MANOVA. A MANOVA statistic can indicate no significant effects but then one way ANOVAs indicate that there are significant differences.

We know why. It makes for interesting discussion. We can talk all we want about experimentwise error and simultaneous effects. The simple reality is that there are instances when a significant overall ANOVA does not result and there are unquestionably significant pairwise differences. For instance, take a categorical variable with five levels. Four of the levels may be very close (yielding small between-group variation) and the fifth might be quite distant (but its effect is not enough to counteract the small between-group variation generated by the other four). Thus, you may find small between-group variation, large within-group variation, and an overall nonsignificant result. The distant group legitimately differs from the others (experimentwise error notwithstanding) but the overall ANOVA result says otherwise.

All this makes for ANOVA being a difficult topic to explain. I personally find regression much more intuitive and easier to present. To help students understand the derivation of the F test, I typically have them compute a sum of squares to demonstrate how the influence of degrees of freedom, the influence of sample size, the influence of within group variation, and the influence of between-group variation influence the overall ANOVA statistic. If explained well, it is not too difficult, and the students (astonishingly) find the exercise enlightening. Here is the little example I use:

For the following 3 distributions compute an F value; the solution follows:

Group A: 2 3 4 4 5 5 6 6 7 8
Group B: 5 6 7 7 8 8 9 9 10 11
Group C: 8 9 10 10 11 11 12 12 13 14

Solution:

- Means for the three groups are 5, 8, and 11, respectively.

- Within-group sums of squares for the three groups are 30, 30, and 30 yielding a total within-groups sum of squares of 90. The within-groups degrees of freedom is 30 – 3 = 27. The within groups mean square is 90/27, or 3.3333.
- The between-groups sum of squares for the three groups are 90, 0, and 90 yielding a total between-groups sum of squares of 180. Degrees of freedom is 2, the number of groups minus 1. The between groups mean square is 180/2 = 90.
- The final F value is the between-groups mean square divided by the within-groups mean square: 90/3.3333 = 27.000, $p < .001$.

We do not spend a lot of time discussing the qualities of different post hoc tests. We goggle briefly at the 22 (!!) different tests that are available. We always (in an introductory class anyway) go with the most liberal one, the LSD (least significant differences) test. We talk in the book a little bit about comparisons of the more popular tests (Scheffe, Bonferroni, Tukey) but I spend little time with it in class. In later chapters of the book these tests are discussed in greater detail.

**Contrasts** is something I always cover with the students. In several of their exercises they are asked to compute particular contrasts; the explanation on page 136 of the book does an adequate job of explaining how contrasts work. You will need to go through some of these on the board, step by step, with students, but once understanding is acquired, most are able to write up contrasts quite adequately.

In the Output section in the book, I feel there is a nice blend of the use of SPSS output, commentary, and definitions to give a clear idea about the meaning of the results for one-way ANOVA. To assist students in writing ANOVA results, I typically give them a generic format chart and suggest that they follow it precisely. Usually students have great difficulty producing a coherent description of ANOVA results. Thus, I suggest that they follow the format strictly and after they've spent a few years writing up results, then they will begin to streamline the rather boxy format suggested below. These write-ups are very difficult to grade unless you specify the strict format. The structured format I use follows:

| |
|---|
| A One-way Analysis of Variance indicated a significant influence of [INDEPENT VARIABLE] on [DEPENDENT VARIABLE], $F(x, xxx) = \_.\_\_$, $p < .xxx$). Post hoc analyses using the [POST HOC METHOD] method with an alpha value of [ALPHA VALUE] found that [LEVEL A] ($M =$ ) was significantly higher than [LEVEL B] ($M =$ ) and that [*repeat for as many pair-wise differences as there are*]. |

# Chapter 13: 2-Way Analysis of Variance

The following chapter (Chapter 14) is so complex (including 3-way ANOVA, covariates, and graphing interactions) that Chapter 13 was purposely kept simple. Within the context of the three ANOVA chapters, in Chapter 13 we extend beyond the simplicity of pairwise comparisons to ask the experimental questions so characteristic of Analysis of Variance:

- Is there a main effect for the first independent variable?
- Is there a main effect for the second independent variable?
- Is there an interactive effect of the two independent variables on the dependent variable?

While ANOVA has its complexities, it also has certain simplicities: For 2-way ANOVA there is exactly one continuous dependent variable and exactly two categorical independent variables. These concepts seem to be learned easily by most students. Then we extend to ANOVA options (see Screen 13.3 on page 145). In the classes I teach, we go exclusively with the Hierarchical method; if covariates are included, we have their influence entered before the effects of the independent variables, and we always include the means and counts so we have visual listing of the mean values in each category.

The example we use in Chapter 13 is the influence of **gender** (2 levels) and **section** (3 levels) on **total** points earned in the class. This provides a simple and straight forward example of a two-way ANOVA. Despite the fact that there are no significant main effects or interactions the example seems to clarify the process for my students. One thing that is emphasized in this chapter is interpretation of output.

We explain that for a main effect with two levels (like gender) there is no reason to conduct post hoc analyses. It is clear from the cell means which group is higher than which. If, however, there is a main effect with a variable that has three or more levels, just like a one-way ANOVA, we may not be certain where the differences lie. When significant main effects occur, I have students systematically conduct a one-way ANOVA with post hoc tests to identify pairwise differences. Once again, my method is a bit rigid and mechanical, but I find that with introductory students they need that structure.

The most difficult part for students is writing up findings. Since the interaction in the Chapter 13 example is not significant, we wait until Chapter 14 before we attempt to write results that include an interaction. But even writing up ANOVA results where there are only main effects, they find challenging. What I have done to assist them and myself (as I go through the nightmare of trying to grade their papers) is to provide a strict structure for write-ups that applies to 2-way or higher order ANOVAs and includes how to write up results if a covariate is included. It is similar to the structure for one-way ANOVA (parts, in fact, are duplicates). It helps them write results and helps me or my TA to grade their papers.

While we present and describe the graph presented in the text, we wait until Chapter 14 before students create graphs of interactions and then attempt to describe the results. Here then is the form that I provide students to assist them in writing the results from 2-way and 3-way ANOVAs:

## How to Write up 2- and 3-way ANOVA Results

A [TYPE OF ANALYSIS] was conducted to determine the influence of [INDEPENDENT VARIABLE] and [INDEPENDENT VARIABLE] and [INDEPENDENT VARIABLE] on [DEPENDENT VARIABLE].

[If there is a covariate] To control for the influence of [NAME OF COVARIATE] it was included as a covariate. [NAME OF COVARIATE] accounted for a significant amount of the overall variance $F$(x, xxx) = _.__, $p$ < .xxx). [*repeat for as many covariates as there are that accounted for a significant portion of the variance*]

There was a main effect found for [INDEPENDENT VARIABLE], $F$(x, xxx) = _.__, $p$ < .xxx). Post hoc analyses using the [POST HOC METHOD] method with an alpha value of [ALPHA VALUE] found that [LEVEL A] ($M$ = ) was significantly higher than [LEVEL B] ($M$ = ) and that [*repeat for as many pairwise differences as there are*] [*repeat this entire section for as many significant main effects as there are*] [*note that if there are only two levels of a variable then post hoc comparisons do not need to be conducted*]

There was [also] a significant [INDEPENDENT VARIABLE] by [INDEPENDENT VARIABLE] interaction $F$(x, xxx) = _.__, $p$ < .xxx). [then describe the interaction based on the graph in general terms]

# Chapter 14: 3-Way Analysis of Variance

Three-way ANOVAs are complex. Armed with that knowledge Paul and I have spent as much time writing Chapter 14 as any chapter in the book. We inch through the material slowly, painstakingly, and attempt to make each prior level clear before we move on to the next level. It is also the longest Output section in the book. I am personally pleased with the result. You can't speed-read the chapter, but anyone with reasonable math aptitude is able to read and to understand. A number of special concerns attend this chapter. What they are and my response to them follows.

First the example is simple and intuitive: the influence of **gender, section**, and **lowup** (lower or upper division student) on **total** points earned. The three independent variables have face validity even to the beginner because it seems essentially reasonable that gender, section, or class standing MAY have an influence on how many points one earns in a class. As such, the example has proved to be a good vehicle to explain three-way ANOVA.

Make clear the seven experimental questions (identified cleanly on pages 150 and 151): With three independent variables there will be three possible main effects (main effects for **gender, section**, and **lowup**), three possible interaction effects (**gender** by **section**, **gender** by **lowup**, and **section** by **lowup**) and one possible 3-way interaction (**gender** by **section** by **lowup**). In the often vague and fuzzy world of ANOVA, it really helps students to see some invariant constants in the process.

The influence of covariates is presented for the first time in this chapter. The example gets the basic idea across effectively. Of course, entire books have been written on analysis of covariance and I am sure such authors would find our brief explanation inadequate. How much time you spend on this topic depends on your class objectives. In my classes we'll spend a half hour on presentation and provide several examples of the use of covariates in different settings. Also, assignments will include one or two ANOVAs that include one or more covariates. How to write up the results of covariates is provided in the model presented on page 38 of this manual.

In the Output section we exert thoughtful effort (pages 160, 161) to provide a more lucid picture of the statistical influence of a covariate. An entire ANOVA output is provided with all statistics included for the analysis *with* the covariate (in bold face to the left) and the identical analysis *without* the covariate (in italics, nonbold, in parentheses, and to the right). We touch on the fact that when a covariate is included (if it has a significant influence) its effect on other variables in the output is typically to decrease the $F$ values and increase the corresponding $p$ values. We provide examples where this happens in the table on page 160 and also show one case where the $F$ actually increases (and *p* diminishes) when the covariate is included. Once again, the time you spend here depends on your class objectives.

The clicks on the computer to conduct an ANOVA are the least of anyone's worries. It is quick, easy (too quick and too easy I think), and in moments students can have pages of bewildering output. I have no comment about the step by step section. Just follow the steps on page 154 (with your own variables of interest) to produce output. Performing these steps always prompts a brief spasm of necessary rhetoric (by the instructor) about the necessity of sound design and careful selection of variables before conducting analysis of variance.

Chapter 14 is where I first have students graph their ANOVA interaction results. This used to be a fairly torturous procedure (involving recreating a small data file that represented cell values from ANOVA results) until SPSS (in version 9.0) allowed multiple line graphs or clustered bar charts to be accessed directly. Now the process described on pages 162-163 is quite intuitive and easy to follow. Instruction on what the graphs mean is up to you.

Writing up interactions is a challenge for just about anyone. I tell my students that I have written results from many interactions and I will still almost never get it quite right the first time. Often several revisions are necessary before a correct and aesthetically pleasing write-up emerges. I do not expect intro students to ever get to the aesthetically pleasing stage. A sample write up for the interaction shown on page 157 might run as follows. Please note that this is not a statistically significant finding, but for practice we write it as if it were:

> While there was little difference for scores earned by men and women in the first two sections, in the third section women scored substantially higher than men.

There are many other ways to say the same thing. It is the rare student indeed who, after seeing some examples, can write up an interaction for a new problem correctly and coherently.

Another interactive effect that we present (and discussed in some detail on pages 157 and 158) is what a graph looks like that demonstrates interactions. We work with the parallel-lines concept: If lines are parallel or near parallel it indicates that interactions are unlikely. The further the lines deviate from parallel, the greater the probability that there is a significant interaction. The standard waivers are inserted here: Graphs can be manipulated by extending or constricting the vertical or horizontal axes or by using only a narrow range of your values to make it appear that there is an interaction when there may be none. Therefore, whether significant interactions occur depends entirely on the ANOVA output, regardless of what the graph looks like. Graphs are designed to assist in understanding interactions, not to identify whether one exists.

For three way interactions, we pretty much throw in the towel for an introductory group. We cleverly arranged for the 3-way interaction in the book example to be non-significant. I explain that they are difficult to understand, difficult to explain, and it requires a thorough understanding of the data to make sense of one. My personal feeling is that it takes a course that focuses almost exclusively on ANOVA before much time will be spent untangling the meaning of three-way interactions.

In my classes I have students write results in correct APA format. Before I created the structured format (the model was presented three pages earlier) it was essentially impossible to grade student papers. Once they had the model, then many were able to describe ANOVA results that were understandable.

An additional issue, uncomfortable for many, is a discussion about **degrees of freedom**. I, like many others, had difficulty explaining the concept. In some frustration I once took all the statistic books that I owned and typed out the glossary definition of degrees of freedom from each of them. There was surprisingly little consensus among the definitions that resulted. On the bright side, we can at least see that they didn't copy their definitions from each other.

I have not yet reached the point of complete comfort on my own treatment of the topic. I typically explain that degrees of freedom is a concept related to the sample size and number of statistical questions that are asked in an analysis. The higher the degrees of freedom, the greater the statistical power. The more questions that you ask of the data, the lower your degrees of freedom become. Thus, in a study that involves many analyses, a large sample is required to provide sufficient degrees of freedom to produce valid results. I personally find that comparing a saturated model in structural equation modeling with models that are less than saturated is the best way for *me* to visualize the concept of the "number of parameters that are free to vary"; but this is not helpful for introductory students. If you have found a good way to present the topic I would welcome an e-mail from you with a description.

# Chapter 15: Simple Linear Regression and Curvilinear Regression

In the very earliest draft of this book this chapter didn't exist. We simply launched (eyes tightly closed, limbs splayed) directly into multiple regression with a happy cry. Statisticians who reviewed the book said it was too much. "You must start with something simpler." So we have. In retrospect, I entirely agree with our critics. Chapter 15 has provided a comfortable transition to the much more complex topic of multiple regression analysis.

We begin our presentation of the material by tying linear regression into the discussion about correlation; indeed, simple linear regression *is* a correlation, that is, multiple $R = r$. We continue with a discussion of phenomena that covary and provide a number of real-world examples. We then draw the connection between the inferences that people make hundreds of times daily and the statistical concept of predicted values. Within this discussion the concept of true versus predicted values emerges and a residual or error term is discussed. We introduce the concept of residuals only briefly, then drop it (as the plot thickens) and refer readers to the chapter on residuals, Chapter 28, if they wish more information.

The data set used to demonstrate linear regression is a fictional file that shows the relationship between pre-test anxiety and actual exam performance. We created the file to provide both a linear relationship (more anxiety yields better exam scores) and a curvilinear relationship (at the high end of the anxiety scale performance begins to drop off). This file provides clear visual evidence (from viewing a scattergram) that both linear and curvilinear relationships exist between anxiety and exam scores. Pages 169 and 170 provide these visual displays.

The topic of predicted values remains the main theme throughout the fairly lengthy (for this book anyway) introduction. We create a linear regression equation then try actual values from subjects in the data set to demonstrate how well (or poorly) the regression equation is able to predict scores. The equation doesn't do very well and this inadequacy leads into the discussion of curvilinear relationships.

Before a discussion of curvilinear trends, however, four key points must be secured:

- The regression command generates a Multiple R value (equal to r when there are only two variables) that measures the strength of relationship between an independent (or predictor) variable and a designated dependent (or criterion) variable.

- Along with the Multiple R, a significance value is produced that identifies the likelihood that the observed value occurred by chance.
- $R^2$ is produced, a statistic that identifies the percent of variation in one variable accounted for by another. For instance, in the linear equation, it is found that 23.8% of the exam score is accounted for by pretest anxiety. We also note that there is no causal ambiguity in that statement.
- The regression output computes the constant and coefficient necessary to create the regression equation.

The discussion that follows about curvilinear trends is sufficiently clear that my students typically find it interesting rather than daunting. Even on the computers they attack with relish the idea of testing for curvilinear trends using the procedure illustrated in steps 4a and 5a (pages 172, 173) that produces the output shown on page 170.

We then create the regression equation that include the curvilinear component, substitute anxiety scores from the same three students in our data set, and discover that the quadratic equation does a superior job of predicting scores than the linear equation was able to do.

For assignments, students complete several linear regression problems. Then we check the same variables for curvilinear trends, and if there is significant curvilinearity of data, we create an additional analysis using both the linear and the curvilinear components. The step-by-step instructions for completing this task are included in Step 5b on page 174. Next students create the regression equations and try out values from three or four subjects to see how well the equation works. The process is very similar to the mode of presentation in the chapter, and students typically find this an easier than average topic to negotiate.

The **divorce.sav** file (see the Exercises section) provides some good examples (using real data) of variables (social support, closeness, income, attributional style, locus) that exert both a linear and a curvilinear influence on the dependent variable, life satisfaction.

# Chapter 16: Multiple Regression Analysis

With thorough attention to the material in Chapter 15, the introduction of multiple regression analysis is greatly facilitated. We begin, once again, with the concept of a regression equation to determine predicted values. We, for the second chapter in a row, shift away from the **grades.sav** file and present the first of the three helping files, **helping1.sav**. This was actually a pilot study (*N* = 81) for several much larger studies that have been conducted since then. For instance the **helping3.sav** file is a data set drawn from more than 1000 subjects (the *N* of 537 represents number of helper-recipient pairs) and has resulted in one publication and has two additional articles (at the time of this writing) under review. This file is used in several of the more complex analyses in later chapters.

An intuitive, simple regression equation is introduced in which the amount of helping is predicted by the amount of **sympathy** the helper feels toward a needy friend. Sympathy is substantially correlated with helping and accounts for about 20% of the variance. Then we continue to suggest that factors other than sympathy may influence how much help is given. Within the example provided, we suggest that **anger** (anger felt by the helper toward his or her friend) and **self-efficacy** (the friend's belief that he or she has the ability to render effective help) also contribute to the helping process. We then demonstrate the regression equation and substitute some numbers from actual subjects to illustrate how well the equations predict the amount of helping.

In the classroom setting I will usually present one or two additional real world examples to underline that most phenomena have multiple causes. A typical illustration is predictors of body weight. Students quickly come up with a string of factors that might influence how much a person weighs: daily caloric intake, height, bone structure, amount of daily exercise, number of fat grams consumed, metabolism. Students seem to grasp the fundamental structure of regression much more easily than they do the fundamentals of ANOVA.

Back to the book example: The regression equation is created and tested to demonstrate its usefulness in predicting values. During this process an unexpected finding is revealed that anger correlates positively with amount of time spent helping. Since one would expect that more anger would result in less help, this provides a good opportunity to identify what an analysis can and cannot do. It can compute valid findings based on certain known parameters. It cannot explain why we get the results we do.

A discussion of the amount of variance explained parallels the presentation in the prior chapter, but this time within the context of multiple regression. We also explain the nature of partial correlation and describe how in a stepwise or forward type of analysis that variables are entered sequentially into the regression equation based on the amount of total variance they explain. When there are no more variables that explain a significant amount of additional variance, then the regression process stops and results are printed. This may entail some careful explanation before students fully grasp it.

There are many classes that spend an entire semester on just regression. Within a few introductory pages, therefore, we cannot begin to cover all relevant points. The section titled "Curvilinear Trends, Model Building, and References," (page 182) raises issues that are well worthy of substantial attention if one wishes students to use the regression procedure successfully. These include:

- Thoughtfully crafted and carefully designed research. This, of course, is a theme which has been reinforced throughout the book.
- Sample size. It isn't as much the number of subjects as the number of subjects compared to the number of analyses you are conducting. I don't know any rule-of-thumb ratios or numbers, but a small sample with many variables would yield an analysis with so few degrees of freedom that significant results would be almost impossible to achieve.
- Examination of data for abnormalities or outliers is worthy of discussion. Use some examples from your own research to illustrate that sometimes forms have to be discarded, or provide examples of outliers in your own data.
- Normal distribution of predictor and criterion variables is worth noting (since normality is the foundation for most forms of data analysis) but need not be over emphasized. The regression equation works quite well with an occasional categorical (ordinal, of course) variable or even a dichotomous variable such as gender. If you attempt to build a case that the dependent (or criterion) variable should be normally distributed, you will get shot down by anyone who has heard of logistic regression or discriminant analysis. Both of them employ a dichotomous variable as the dependent variable. In general, mathematicians have found that the regression equation often works quite well even when assumptions of normality are violated.
- The issue of linear dependency is no trivial after thought. I spend a good deal of time addressing this issue in my classes. The first concern is to eliminate predictor variables that are numerically dependent (such as quiz scores and total points, or a measure of apprehensiveness that is a mathematical composite of other variables in the data set such as tenseness, suspiciousness, and anxiety). Next is to look carefully at variables that that may be conceptually similar, such as anxiety and tension. An initial correlation matrix can often help avoid potential difficulties by noticing variables in the matrix that are highly correlated with each other. An example from the **divorce.sav** file: The attributional style questionnaire (ASQ, optimistic or pessimistic attributional style) and locus of control were included in initial analyses. After it was observed that ASQ and locus were highly correlated,

we discovered that the ASQ incorporates a locus component into its measure. Locus was subsequently dropped from the analysis.

The Step by Step section is a bit more involved than the equivalent section from Chapter 15. Certain ideas should be addressed: For instance, the different methods of entering variables into the regression equation should be considered. The forward and stepwise methods are most frequently used, and there is not time or space in an introductory class to try and compare several of them. Students, however, should be aware that different options exist.

The **Plots** option deals largely with plots of residuals and is not covered except in more advanced classes. The default statistics are satisfactory for introductory students. The **Save** option provides for some interesting comparisons. If, for instance, you save the unstandardized predicted values, you can place them in the column next to the actual values of the dependent variable and get a first hand look at how well the regression equation does at predicting the criterion variable.

Moving now to the Output section, an additional area of concern is distinguishing between B values and beta values. Explain that the B values are the correct values to use in creating the regression equation. However, these values are not comparable with each other because the metrics on which they are based might vary widely. The beta values, on the other hand, are based on standardized scores and thus vary strictly between +1 and –1. Betas can be compared directly to identify the magnitude of influence of a particular predictor variable on the criterion variable. Also, the sign (+ or **-**) of beta values has unique significance. A positive sign indicates that more of the predictor variable results in more of the criterion variable. A negative beta indicates that more of the predictor results in less of the criterion variable. It is helpful to use a number of examples to illustrate this.

For assignments, I typically provide several data sets (available at **www.abacon.com/george**) that are particularly conducive to regression analysis. There are a number of different regression outputs in the Exercises section of this manual. Then I have students run the procedures, and, for written output, they create a chart that lists the Multiple R, the $R^2$, and then record variables that significantly predict the criterion variable, listed in order of magnitude of the beta values with the beta values included in parentheses for each variable.

Some defined terms in the last two or three pages are also of value. Particularly useful are the meaning of **beta**, **partial correlation**, **minimum tolerance** (another tool to help avoid linear dependency in your analyses), and **r-square change**.

# Section III

# Exercises

# Exercises

This section provides a wealth of resources for the instructor attempting to find meaningful exercises for his or her students. With only minimal knowledge of a particular data set, it might take hours of trying different analyses before you come up with the results that will illustrate the points or hone the skills that you desire.

All exercises that follow are based on data included on two separate data files. Both are included on the enclosed data disk. First the **divorce.sav** file will be described and presented, then exercises based on that file, followed by a description of the **helping3.sav** data file, then exercises based on that file. Types of exercises for each file will be presented in the following order:

- chi-square analyses,
- t-tests,
- 1-way ANOVAs
- 2- and 3- way ANOVAs,
- curve estimation, and
- regressions.

Analyses such as frequencies, descriptives, and the means procedure are so self evident that they do not need examples in a manual of this sort to illustrate them. Correlations also, although one of the most frequently used types of analyses in research, are so simple to compute that to present examples here would be a waste of space. The procedures listed above would require a lot of time for an instructor to find examples that would produce significant results. This section has, therefore, been created to provide examples that can become exercises for your students. We now turn our attention to a description of the first file, **divorce.sav**.

## The DIVORCE.SAV file (available at **www.abacon.com/george**)

**divorce.sav**: This is a file of 229 divorced individuals recruited from communities in Central Alberta. The objective of researchers was to identify cognitive or interpersonal factors that assisted in recovery from divorce. Key variables employed in the study included:

<u>Dependent variables</u>

**lsatisfy**: A measure of life satisfaction based on weighted averages of satisfaction in 12 different areas of life functioning. This is scored on a 1 (low satisfaction) to 7 (high satisfaction) scale.

**trauma**: A measure of the trauma experienced during the divorce recovery phase based on the mean of 16 different potentially traumatic events, scored on a 1 (low trauma) to 7 (high trauma) scale.

Demographics

**sex**: [women(1), men(2)]
**age**: (ranges from 23 to 76)
**sep**: (years separated accurate to one decimal)
**mar**: (years married prior to separation, accurate to one decimal)
**status**: present marital status [married(1), separated(2), divorced(3), cohabiting(4)]
**eth**: ethnicity [White(1), Hispanic(2), Black(3), Asian(4), other or DTS(5)]
**school**: [1-11yr(1), 12yr(2), 13yr(3), 14yr(4), 15yr(5), 16yr(6), 17yr(7), 18yr(8), 19+(9)]
**childneg**: number of children negotiated in divorce proceedings
**childcst**: number of children presently in custody
**income**: [DTS(0), <10,000(1), 10-20(2), 20-30(3), 30-40(4), 40-50(5), 50+(6)]

Key independents

**cogcope**: amount of cognitive coping during recovery [little(1) to much(7)]
**behcope**: amount of behavioral coping during recovery [little(1) to much(7)]
**avoicope**: amount of avoidant coping during recovery [little(1) to much(7)]
**iq**: intelligence or ability at abstract thinking [low(1) to high(12)]
**close**: amount of physical (non-sexual) closeness experienced [little(1) to much(7)]
**locus**: locus of control [external locus(1) to internal locus(10)]
**asq**: attributional style questionnaire [pessimistic style(≈-7) to optimistic style(≈+9)]
**socsupp**: amount of social support experienced [little(1) to much(7)]
**spiritua**: level of personal spirituality [low(1) to high(7)]

**Chi-Square Analysis**: **sex** (2 levels) by level of **income**: (5 levels). Note that the level 0 (decline to state) and level 6 (over $50,000) of **income** have been removed prior to the analysis due to so few subjects appearing in either category.

```
SEX  by  INCOME  current family income

                    INCOME                                Page 1 of 1
            Count  |
           Exp Val |<10,000  10-20,00 20-30,00 30-40,00 40-50,00
                   |         0        0        0        0        Row
                   |     1  |     2  |     3  |     4  |     5  | Total
SEX        --------+--------+--------+--------+--------+--------+
                1  |    23  |    35  |    23  |    17  |    16  |   114
  female           |  18.2  |  26.5  |  28.6  |  20.8  |  19.8  | 52.1%
                   +--------+--------+--------+--------+--------+
                2  |    12  |    16  |    32  |    23  |    22  |   105
  male             |  16.8  |  24.5  |  26.4  |  19.2  |  18.2  | 47.9%
                   +--------+--------+--------+--------+--------+
            Column      35       51       55       40       38      219
             Total   16.0%    23.3%    25.1%    18.3%    17.4%   100.0%

    Chi-Square                 Value          DF              Significance
------------------          -----------      ----             ------------
Pearson                        13.50862         4                   .00904
Likelihood Ratio               13.73259         4                   .00820
Mantel-Haenszel test for        9.00225         1                   .00270
      linear association

Minimum Expected Frequency -   16.781

                                                               Approximate
Statistic                     Value        ASE1      Val/ASE0   Significance
----------------            ---------    --------    --------   ------------
Phi                           .24836                                .00904 *1
Cramer's V                    .24836                                .00904 *1
```

**Chi-Square Analysis**: marital **status** (4 levels) by level of **income**: (5 levels). Note that the level 0 (decline to state) and level 6 (over $50,000) of **income** have been removed prior to the analysis due to so few subjects appearing in either category.

```
STATUS  current marital status  by  INCOME  current family income

                    INCOME                                 Page 1 of 1
            Count  |
           Exp Val |<10,000  10-20,00 20-30,00 30-40,00 40-50,00
                   |          0        0        0        0         Row
                   |     1  |     2  |     3  |     4  |     5  | Total
STATUS     --------+--------+--------+--------+--------+--------+
                1  |     3  |     9  |    14  |    12  |    15  |    53
  marred           |   8.5  |  12.3  |  13.3  |   9.7  |   9.2  | 24.2%
                   +--------+--------+--------+--------+--------+
                2  |     5  |    10  |     8  |     5  |     4  |    32
  separated        |   5.1  |   7.5  |   8.0  |   5.8  |   5.6  | 14.6%
                   +--------+--------+--------+--------+--------+
                3  |    25  |    22  |    29  |    19  |    14  |   109
  divorced         |  17.4  |  25.4  |  27.4  |  19.9  |  18.9  | 49.8%
                   +--------+--------+--------+--------+--------+
                4  |     2  |    10  |     4  |     4  |     5  |    25
  cohabit          |   4.0  |   5.8  |   6.3  |   4.6  |   4.3  | 11.4%
                   +--------+--------+--------+--------+--------+
            Column      35       51       55       40       38      219
             Total   16.0%    23.3%    25.1%    18.3%    17.4%   100.0%

 Chi-Square                  Value           DF              Significance
---------------            -----------      ----             ------------
Pearson                      20.27895        12                  .06199
Likelihood Ratio             20.45456        12                  .05896
Mantel-Haenszel test for      6.64273         1                  .00996
      linear association
Minimum Expected Frequency -    3.995
Cells with Expected Frequency < 5 -     3 OF    20 ( 15.0%)
                                                                  Approximate
Statistic                    Value        ASE1      Val/ASE0   Significance
---------------            ---------    --------    --------   ------------
Phi                          .30430                                .06199 *1
Cramer's V                   .17569                                .06199 *1
```

**Independent-samples t-test**: Are there gender differences on the length of marriage prior to separation?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
MAR  years married prior to separation
fema                          119      10.4067       7.308         .670
male                          110       8.4900       6.720         .641
-----------------------------------------------------------------------
         Mean Difference = 1.9167
         Levene's Test for Equality of Variances: F= .311   P= .578

      t-test for Equality of Means                                     95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.06      227          .040           .930         (.084, 3.749)
Unequal        2.07   226.99          .040           .927         (.090, 3.743)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the number of children currently in custody?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
CHILDCST
fema                          119       1.4874       1.227         .113
male                          110        .5091        .875         .083
-----------------------------------------------------------------------
         Mean Difference = .9783
         Levene's Test for Equality of Variances: F= 16.603 P= .000

      t-test for Equality of Means                                     95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          6.89      227          .000           .142         (.699, 1.258)
Unequal        6.98   213.57          .000           .140         (.702, 1.254)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on level of income [coded on a low(1) to high(7) scale]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
INCOME  current family income
fema                          119       2.7059       1.434         .131
male                          110       3.1636       1.411         .135
-----------------------------------------------------------------------
         Mean Difference = -.4578
         Levene's Test for Equality of Variances: F= .512   P= .475

      t-test for Equality of Means                                     95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -2.43      227          .016           .188        (-.829, -.087)
Unequal       -2.43   226.11          .016           .188        (-.828, -.087)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the amount of cognitive coping used in divorce recovery [coded on a low(1) to high(7) scale]?

```
                              Number
Variable                     of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------------
COGCOPE  cognitive-active coping
fema                           119        4.5309        .884          .081
male                           110        4.2751        .974          .093
-----------------------------------------------------------------------------
         Mean Difference = .2558
         Levene's Test for Equality of Variances: F= .031   P= .859

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.08      227           .038           .123          (.014, .498)
Unequal        2.08   220.21           .039           .123          (.013, .499)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the amount of avoidant coping used in divorce recovery [coded on a low(1) to high(7) scale]?

```
                              Number
Variable                     of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------------
AVOICOP:  Avoidant coping
fema                           119        2.5473        .844          .077
male                           110        2.9216        .956          .091
-----------------------------------------------------------------------------
         Mean Difference = -.3743
         Levene's Test for Equality of Variances: F= 3.938  P= .048

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -3.15      227           .002           .119        (-.609, -.140)
Unequal       -3.13   217.99           .002           .120        (-.610, -.139)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the amount non-sexual physical closeness experienced [coded on a low(1) to high(7) scale]?

```
                              Number
Variable                     of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------------
CLOSE  amount of physical closeness
fema                           119        3.5067        .943          .086
male                           110        3.2273        .931          .089
-----------------------------------------------------------------------------
         Mean Difference = .2794
         Levene's Test for Equality of Variances: F= .948   P= .331

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.26      227           .025           .124          (.035, .524)
Unequal        2.26   226.01           .025           .124          (.035, .523)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the amount of social support experienced [coded on a little(1) to much(7) scale]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
SOCSUPP  social support
fema                          119        3.7661       1.046         .096
male                          110        3.4272        .839         .080
-----------------------------------------------------------------------
         Mean Difference = .3388
         Levene's Test for Equality of Variances: F= 7.811  P= .006

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.69      227          .008           .126          (.091, .587)
Unequal        2.71   222.70          .007           .125          (.093, .585)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on attributional style [coded on a pessimistic attributional style(-6) to optimistic attributional style(+9) scale]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
ASQ:  Attributional style
fema                          119        3.4386       2.741         .251
male                          110        2.6186       2.688         .256
-----------------------------------------------------------------------
         Mean Difference = .8199
         Levene's Test for Equality of Variances: F= .124   P= .725

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.28      227          .023           .359         (.112, 1.528)
Unequal        2.28   226.20          .023           .359         (.113, 1.527)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there gender differences on the experience of personal spirituality [coded on a not spiritual(1) to very spiritual(7) scale]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
SPIRITUA:  Level of personal spirituality
fema                          119        4.8020       1.077         .099
male                          110        4.1383       1.293         .123
-----------------------------------------------------------------------
         Mean Difference = .6638
         Levene's Test for Equality of Variances: F= 4.546  P= .034

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          4.23      227          .000           .157          (.355, .973)
Unequal        4.20   212.78          .000           .158          (.352, .975)
```

**One-way ANOVA**: Does one's marital **status** (4 levels) have an influence on their level of **income** [coded on a low(1) to high(7) scale]?

```
- - - - -  O N E W A Y  - - - - -

      Variable  INCOME     current family income
   By Variable  STATUS     current marital status

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups             3       25.3196        8.4399       4.2538  .0060
Within Groups            225      446.4183        1.9841
Total                    228      471.7380

Multiple Range Tests:  LSD test with significance level .05

The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .9960 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.79
   (*) Indicates significant differences which are shown in the lower triangle

                          S D C M
                          e i o a
                          p v h r
                          a o a r
                          r r b i
     Mean      STATUS
     2.6176    Separ
     2.7304    Divor
     3.0000    Cohab
     3.4909    Marri      * *
```

**One-way ANOVA**: Is one's marital status (4 levels) associated with the number of years of schooling completed [coded on a <11yr(1) to >19 yr(10) scale]?

```
- - - - -  O N E W A Y  - - - - -

      Variable  SCHOOL     number of years of schooling
   By Variable  STATUS     current marital status

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups             3       48.4137       16.1379       2.4914  .0610
Within Groups            225     1457.4291        6.4775
Total                    228     1505.8428

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 1.7996 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.79
   (*) Indicates significant differences which are shown in the lower triangle

                          S M C D
                          e a o i
                          p r h v
                          a r a o
                          r i b r
     Mean      STATUS
     3.2941    Separ
     3.6545    Marri
     3.8400    Cohab
     4.4609    Divor      *
```

**One-way ANOVA**: Is one's marital status (4 levels) associated with chronological age [range, 23-76]?

```
                       - - - - -  O N E W A Y  - - - - -

      Variable  AGE
   By Variable  STATUS     current marital status

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups             3     1569.7781      523.2594       5.6896  .0009
Within Groups            225    20692.9118       91.9685
Total                    228    22262.6900

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 6.7812 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.79
   (*) Indicates significant differences which are shown in the lower triangle

                          C S D M
                          o e i a
                          h p v r
                          a a o r
                          b r r i
     Mean      STATUS
    36.6400    Cohab
    39.6765    Separ
    41.9826    Divor      *
    45.4909    Marri      * * *
```

**One-way ANOVA**: Is one's marital status (4 levels) associated with the amount of behavioral coping that is practiced [coded on a little(1) to much(7) scale]?

```
- - - - -  O N E W A Y  - - - - -

      Variable  BEHCOPE
   By Variable  STATUS     current marital status

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups             3       11.3115        3.7705       3.4029  .0185
Within Groups            225      249.3047        1.1080
Total                    228      260.6162

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .7443 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.79
   (*) Indicates significant differences which are shown in the lower triangle

                          C M D S
                          o a i e
                          h r v p
                          a r o a
                          b i r r
     Mean      STATUS
     4.0444    Cohab
     4.0929    Marri
     4.5382    Divor      * *
     4.5588    Separ        *
```

**One-way ANOVA**: Is one's marital status (4 levels) associated with one's level of personal spirituality [coded on a low(1) to high(7) scale]?

```
- - - - -  O N E W A Y  - - - - -

      Variable  SPIRITUA
   By Variable  STATUS     current marital status

                                  Sum of         Mean             F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups              3      17.4872        5.8291       4.0150  .0083
Within Groups             225     326.6600        1.4518
Total                     228     344.1472

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .8520 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.79
   (*) Indicates significant differences which are shown in the lower triangle

                          C S D M
                          o e i a
                          h p v r
                          a a o r
                          b r r i
     Mean      STATUS

     3.7200    Cohab
     4.4982    Separ      *
     4.5425    Divor      *
     4.6970    Marri      *
```

**Three-way ANOVA with 2 ´ main effects and 1 ´ 2-way interaction**: What influence does **sex** (2 levels), marital **status** (4 levels), and **income** (7 levels) have on one's personal spirituality [coded on a low(1) to high(7) scale]?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                SPIRITUA
           by   SEX
                STATUS   current marital status
                INCOME   current family income

                HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 44.761 | 8 | 5.595 | 4.356 | .000 |
| **SEX** | **24.578** | **1** | **24.578** | **19.136** | **.000** |
| **STATUS** | **13.942** | **3** | **4.647** | **3.618** | **.014** |
| INCOME | 6.240 | 4 | 1.560 | 1.215 | .306 |
| 2-Way Interactions | 41.878 | 19 | 2.204 | 1.716 | .037 |
| SEX STATUS | 5.580 | 3 | 1.860 | 1.448 | .230 |
| SEX INCOME | 4.268 | 4 | 1.067 | .831 | .507 |
| **STATUS INCOME** | **30.139** | **12** | **2.512** | **1.955** | **.031** |
| 3-Way Interactions | 14.129 | 8 | 1.766 | 1.375 | .210 |
| SEX STATUS INCOME | 14.129 | 8 | 1.766 | 1.375 | .210 |
| Explained | 100.767 | 35 | 2.879 | 2.242 | .000 |
| Residual | 235.043 | 183 | 1.284 | | |
| Total | 335.810 | 218 | 1.540 | | |

**Three-way ANOVA with 2 ´ main effects and 1 ´ 3-way interaction**: What influence does **sex** (2 levels), marital **status** (4 levels), and **income** (7 levels) have one the amount of non-sexual closeness experienced [coded on a low(1) to high(7) scale]?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                CLOSE    amount of physical closeness
           by   SEX
                STATUS   current marital status
                INCOME   current family income

                HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 17.593 | 8 | 2.199 | 2.678 | .008 |
| **SEX** | **4.156** | **1** | **4.156** | **5.060** | **.026** |
| STATUS | 2.654 | 3 | .885 | 1.077 | .360 |
| **INCOME** | **10.783** | **4** | **2.696** | **3.283** | **.013** |
| 2-Way Interactions | 12.046 | 19 | .634 | .772 | .738 |
| SEX STATUS | 3.259 | 3 | 1.086 | 1.323 | .268 |
| SEX INCOME | 4.817 | 4 | 1.204 | 1.466 | .214 |
| STATUS INCOME | 7.621 | 12 | .635 | .773 | .677 |
| 3-Way Interactions | 13.779 | 8 | 1.722 | 2.097 | .038 |
| **SEX STATUS INCOME** | **13.779** | **8** | **1.722** | **2.097** | **.038** |
| Explained | 43.419 | 35 | 1.241 | 1.511 | .044 |
| Residual | 150.292 | 183 | .821 | | |
| Total | 193.711 | 218 | .889 | | |

**Three-way ANOVA with 2 ´ main effects and 1 ´ 3-way interaction**: What influence does **sex** (2 levels), marital **status** (4 levels), and **income** (7 levels) have on intelligence [coded on a low(1) to high(12) scale]?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 IQ
            by   SEX
                 STATUS   current marital status
                 INCOME   current family income

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 85.433 | 8 | 10.679 | 2.790 | .006 |
| SEX | .739 | 1 | .739 | .193 | .661 |
| **STATUS** | **31.247** | **3** | **10.416** | **2.721** | **.046** |
| **INCOME** | **53.447** | **4** | **13.362** | **3.491** | **.009** |
| 2-Way Interactions | 83.633 | 19 | 4.402 | 1.150 | .306 |
| SEX STATUS | 7.256 | 3 | 2.419 | .632 | .595 |
| SEX INCOME | 7.868 | 4 | 1.967 | .514 | .726 |
| STATUS INCOME | 69.227 | 12 | 5.769 | 1.507 | .125 |
| 3-Way Interactions | 65.210 | 8 | 8.151 | 2.129 | .035 |
| **SEX STATUS INCOME** | **65.210** | **8** | **8.151** | **2.129** | **.035** |
| Explained | 234.276 | 35 | 6.694 | 1.749 | .010 |
| Residual | 700.500 | 183 | 3.828 | | |
| Total | 934.776 | 218 | 4.288 | | |

**Three-way ANOVA with 1 ´ main effect and 1 ´ 2-way interaction**: What influence does **sex** (2 levels), marital **status** (4 levels), and **income** (7 levels) have on the amount of avoidant coping practiced [coded on a low(1) to high(7) scale]?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 AVOICOP
            by   SEX
                 STATUS   current marital status
                 INCOME   current family income

                 HIERARCHICAL sums of squares
```

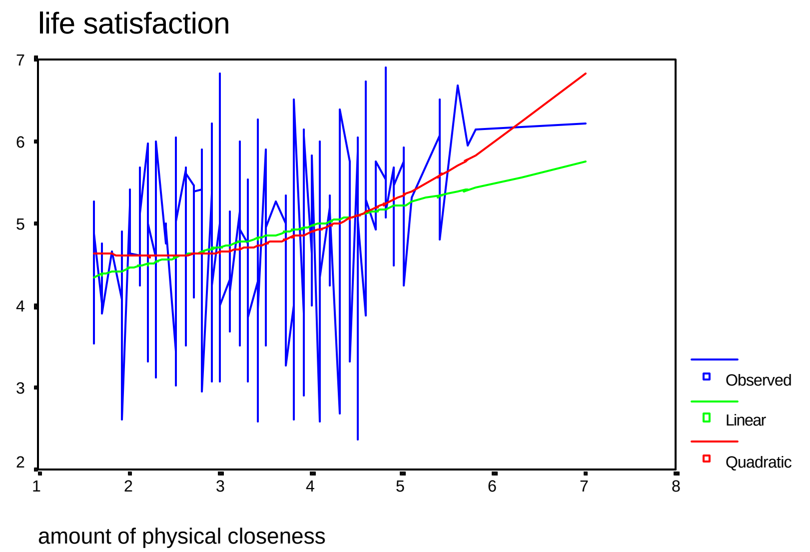
| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 14.620 | 8 | 1.828 | 2.320 | .021 |
| **SEX** | **8.198** | **1** | **8.198** | **10.407** | **.001** |
| STATUS | 4.445 | 3 | 1.482 | 1.881 | .134 |
| INCOME | 1.978 | 4 | .494 | .628 | .643 |
| 2-Way Interactions | 27.098 | 19 | 1.426 | 1.811 | .024 |
| SEX STATUS | .663 | 3 | .221 | .280 | .840 |
| SEX INCOME | 4.364 | 4 | 1.091 | 1.385 | .241 |
| **STATUS INCOME** | **21.467** | **12** | **1.789** | **2.271** | **.010** |
| 3-Way Interactions | 3.342 | 8 | .418 | .530 | .833 |
| SEX STATUS INCOME | 3.342 | 8 | .418 | .530 | .833 |
| Explained | 45.061 | 35 | 1.287 | 1.634 | .021 |
| Residual | 144.147 | 183 | .788 | | |
| Total | 189.208 | 218 | .868 | | |

**Multiple Regression Analysis**: Criterion variable: Life satisfaction (**lsatisfy**) with predictors of **income**, avoidant coping (**avoicope**), social support (**socsupp**), curvilinear influence of social support (**socsupp2**), closeness (**close**), curvilinear influence of closeness (**close2**), attributional style (**asq**), curvilinear influence of attributional style (**asq**), gender (**sex**), **age**, length of separation (**sep**), years married prior to separation (**mar**), amount of education (**school**), cognitive coping (**cogcope**), behavioral coping (**behcope**), intelligence (**iq**), level of spirituality (**spiritua**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   LSATISY   life satisfaction

Variable(s) Entered on Step Number
   5..    ASQ2

Multiple R           .53411
R Square             .28527
Adjusted R Square    .26849
Standard Error       .76462

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           5            49.70381          9.94076
Residual           213           124.52989           .58465
F =      17.00300       Signif F =  .0000

------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
INCOME          .181216     .039199    .268637     4.623  .0000
AVOICOP        -.197333     .057160   -.205638    -3.452  .0007
SOCSUPP         .197159     .059579    .211667     3.309  .0011
CLOSE2          .024225     .008152    .185128     2.971  .0033
ASQ2            .005001     .002434    .124664     2.055  .0411
(Constant)     3.725877     .300442               12.401  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
SEX          .029057  .032137    .806335      .468  .6401
AGE          .021013  .024187    .819084      .352  .7250
SEP          .005574  .006500    .819494      .095  .9247
MAR          .077389  .089394    .816446     1.307  .1927
SCHOOL      -.084724 -.094504    .799803    -1.382  .1684
COGCOPE      .007203  .008034    .809006      .117  .9070
BEHCOPE     -.053882 -.060262    .781145     -.879  .3804
IQ          -.017934 -.020721    .819885     -.302  .7631
CLOSE       -.444195 -.098248    .034966    -1.437  .1521
ASQ         -.015735 -.009917    .283935     -.144  .8853
SPIRITUA    -.002106 -.002264    .740952     -.033  .9737
SOCSUP2     -.217159 -.041230    .025764     -.601  .5486

End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: Criterion variable: Severity of trauma experienced during recovery **(trauma)** with predictors of **income**, avoidant coping (**avoicope**), social support (**socsupp**), curvilinear influence of social support (**socsupp2**), closeness (**close**), curvilinear influence of closeness (**close2**), attributional style (**asq**), curvilinear influence of attributional style (**asq2**), gender (**sex**), **age**, length of separation (**sep**), years married prior to separation (**mar**), amount of education (**school**), cognitive coping (**cogcope**), behavioral coping (**behcope**), intelligence (**iq**), level of spirituality (**spiritua**),

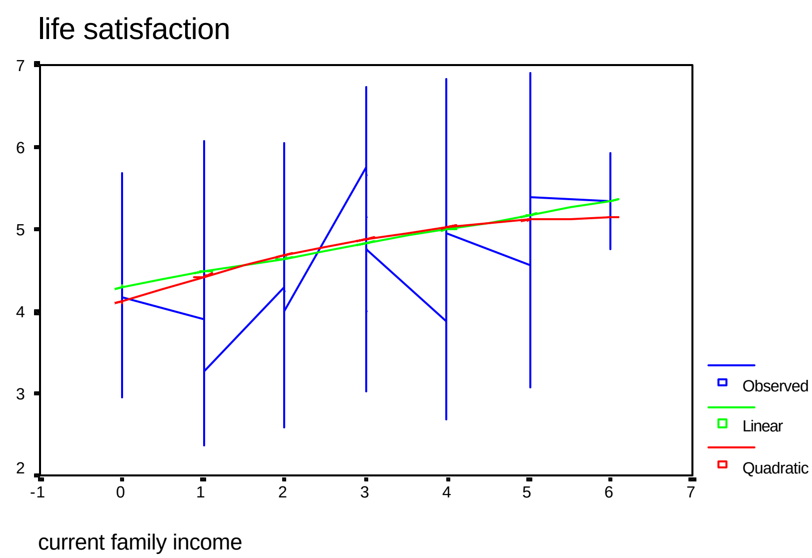
```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   TRAUMA

Variable(s) Entered on Step Number   5..    CLOSE2

Multiple R           .51581
R Square             .26606
Adjusted R Square    .24883
Standard Error       .92439

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           5            65.97867         13.19573
Residual           213           182.00702           .85449
F =      15.44276       Signif F =  .0000

------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
SEP             .023314     .009643    .148126     2.418  .0165
INCOME        -.119309     .047893   -.148250    -2.491  .0135
BEHCOPE         .276516     .060877    .281616     4.542  .0000
AVOICOP         .428601     .067597    .374377     6.341  .0000
CLOSE2        -.018951     .009351   -.121393    -2.027  .0440
(Constant)     1.778256     .367491                 4.839  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
SEX         -.063733 -.069174    .864587    -1.010  .3138
AGE         -.029376 -.027124    .611688     -.395  .6932
MAR         -.021167 -.024163    .895589     -.352  .7252
SCHOOL       .006583  .007312    .896310      .106  .9153
COGCOPE      .019742  .018976    .669859      .276  .7826
IQ          -.033216 -.037644    .889655     -.548  .5839
CLOSE        .491906  .107373    .034969     1.572  .1173
SOCSUPP     -.001704 -.001779    .800428     -.026  .9794
ASQ          .056492  .062959    .875201      .919  .3594
ASQ2         .017718  .019926    .885552      .290  .7720
SOCSUP2     -.014129 -.014629    .786879     -.213  .8315

End Block Number   1   PIN =     .060 Limits reached.
```




The linear and curvilinear influence of physical, non-sexual closeness (**close**), on life satisfaction (**lsatisfy**).

Independent: CLOSE

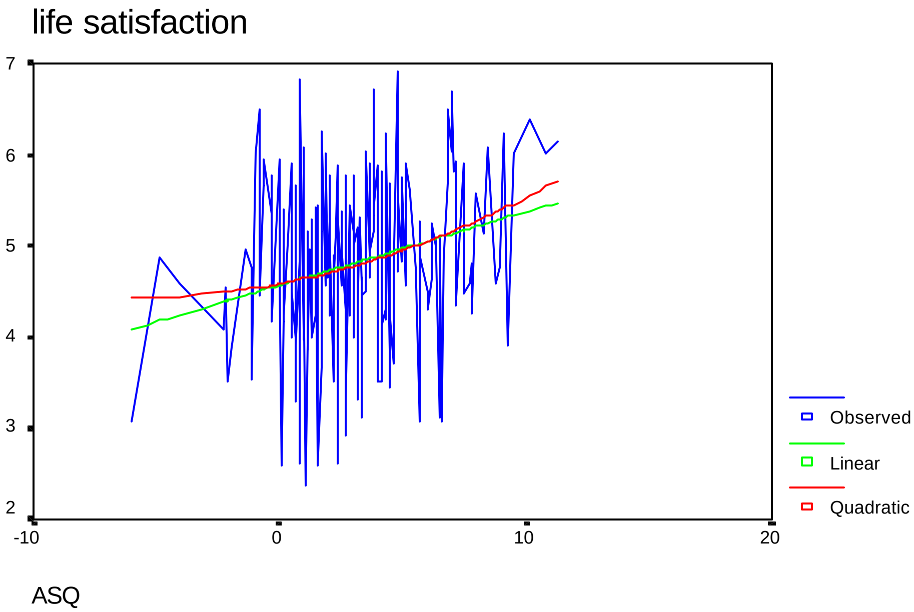
| Dependent | Mth | Rsq | d.f. | F | Sigf | b0 | b1 | b2 |
|---|---|---|---|---|---|---|---|---|
| LSATISY | LIN | .075 | 227 | 18.38 | .000 | 3.9405 | .2593 | |
| LSATISY | QUA | .095 | 226 | 11.87 | .000 | 5.0966 | -.4422 | .0987 |




The linear and curvilinear influence of **income**, on life satisfaction (**lsatisfy**).

Independent: INCOME

| Dependent | Mth | Rsq | d.f. | F | Sigf | b0 | b1 | b2 |
|---|---|---|---|---|---|---|---|---|
| LSATISY | LIN | .079 | 227 | 19.37 | .000 | 4.3042 | .1746 | |
| LSATISY | QUA | .083 | 226 | 10.24 | .000 | 4.1243 | .3356 | -.0274 |




The linear and curvilinear influence of optimistic or pessimistic attributional style (**asq**) on life satisfaction (**lsatisfy**).

Independent: ASQ

| Dependent | Mth | Rsq | d.f. | F | Sigf | b0 | b1 | b2 |
|---|---|---|---|---|---|---|---|---|
| LSATISY | LIN | .060 | 227 | 14.57 | .000 | 4.5707 | .0803 | |
| LSATISY | QUA | .064 | 226 | 7.69 | .001 | 4.5872 | .0510 | .0043 |

## The HELPING3.SAV file (available at **www.abacon.com/george**)

**helping3.sav**: This is a file that can demonstrate any statistical procedure that exists. In articles that have been accepted for publication (or are currently under review), structural equation modeling was the major form of analysis, but every other form of analysis preceded it. It is a study of helping among friends. The N of 537 represents the number of helpers (drawn from a community sample in the Los Angeles area) and 467 help recipients also responded. A procedure described under the missing values section in Chapter 4 dealt with replacing missing values with predicted values from regression equations. This procedure was enacted for the 70 helper forms that did not have equivalent recipient forms. This represented only 2.9% of total data replaced (the recipient forms were much less extensive than the helper forms), well within the limits for psychometric validity. It raised some eyebrows among reviewers, but after explanation, they were satisfied.

The goal of the study was to create a theoretical model of helping among friends. Specifically we were seeking to find out what factors significantly influenced three different dependent help variables. Key variables and demographics used in the study are listed below.

<u>Dependent variables</u>:

**thelplnz**: Time spent helping: [the initial “t” (total) indicates that it is a combination of helper and recipient responses, the latter “lnz” indicates a natural log and z-score transformation]. The z score gives it a mean of 0 and a standard deviation of 1.0 and a range of approximately –3 to +3.

**tqualitz**: Help quality: [the initial “t” (total) indicates that it is a combination of helper and recipient responses, the latter z indicates a z-score transformation]. The z score gives it a mean of 0 and a standard deviation of 1.0 and a range of approximately –3 to +3.

**tothelp**: Total help. The time helping and help quality measures weighted equally. Mean is -.01, standard deviation about .8 and a range –3 to +3 range.

Please don’t confuse your students with the transformational garbage. Just “time helping”, “help quality” and the “combination” should be sufficient. Good opportunity to talk about linear dependency. Time helping and help quality are NOT linearly dependent and any forms of analyses may be used comparing them. Total help, on the other hand, is a mathematical composite of the other two; absolutely linearly dependent.

<u>Sub Categories of the Dependent Variables</u>

**cathelp**: dichotomous variable [less than average(1), more than average(2)]
**empahelp**: amount of time (in hours) spent giving empathic help
**insthelp**: amount of time (in hours) spent giving instrumental help (doing things)
**infhelp**: amount of time (in hours) spent giving informational help

Demographics and other categorical variables:

**gender**: [women(1), men(2)]
**age**: range, 17-89, mean 31.3
**occupat**: [professional(1), service(2), blue-collar(3), unemployed(4), student(5), DTS(6)]
**marital**: marital status [married(1), single(2), DTS(3)]
**school**: [1-8yr(1), 9-11yr(2), 12yr(3), 13-14yr(4), 15-16yr(5), 17-18yr(6), 19+yr(7)]
**ethnic**: ethnicity [Caucasian(1), Black(2), Hispanic(3), Asian(4), other or DTS(5)]
**children**: Number of children living in the household
**problem**: problem type [goal disruptive(1), relational(2), illness(3), catastrophic(4)]
**income**: [<15,000(1), 15-25,000(2), 25-50,000(3), 50,000+(4), DTS(5)]

Good opportunity to use the select cases procedures to eliminate the DTS crowd for some on your Chi-square or ANOVA analyses. However, sometimes the DTS group is interesting in its own right. There are predictable characteristics of individuals who endorse DTS frequently such as more cynical, more suspicious, less spiritual.

Key predictor variables

All variables are the helper's rating of each construct

**hclose**: closeness of the relationship [distant(1) to close(7)]
**hseveret**: severity of the problem [mild(1) to severe(7)]
**angert**: anger felt toward the friend [little(1) to much(7)]
**controt**: controllability; fault, responsibility [not at fault(1) to entirely at fault(7)]
**sympathi**: sympathy felt toward the friend [little(1) to much(7)]
**worry**: worry experienced about the friend's problem [little(1) to much(7)]
**obligat**: feelings of obligation toward the friend [little(1) to much(7)]
**hcopet**: perception of how well the friend is coping [coping poorly(1) to well(7)]
**effict**: helpers belief of self-efficacy (has the ability to help) [low(1) to high(7)]
**empathyt**: helper's empathic tendency [low empathy(1) to high empathy(7)]

There are more variables, of course, and may be understood and used by viewing the variable labels in the data file.

**Chi-Square Analysis**: **gender** (2 levels) by type of problem experienced (**problem**, 4 levels).

```
GENDER  by  PROBLEM  TYPE OF PROBLEM EXPERIENCED

                    PROBLEM                        Page 1 of 1
            Count  |
           Exp Val |GOAL DIS RELATION ILLNESS  CATASTRO
                   |RUPTIVE  AL BREAK          PHIC       Row
                   |     1  |     2  |     3  |     4  | Total
GENDER     --------+--------+--------+--------+--------+
               1   |   119  |   148  |    51  |     7  |   325
  FEMALE           | 138.6  | 126.5  |  52.0  |   7.9  | 60.5%
                   +--------+--------+--------+--------+
               2   |   110  |    61  |    35  |     6  |   212
  MALE             |  90.4  |  82.5  |  34.0  |   5.1  | 39.5%
                   +--------+--------+--------+--------+
            Column     229      209       86       13      537
             Total   42.6%    38.9%    16.0%     2.4%   100.0%

      Chi-Square                 Value           DF              Significance
Pearson                        16.57838           3                 .00086
Likelihood Ratio               16.80914           3                 .00077
Mantel-Haenszel test for        3.45666           1                 .06300
      linear association
Minimum Expected Frequency -    5.132
                                                                 Approximate
     Statistic                   Value       ASE1      Val/ASE0   Significance
--------------------            ---------    --------    --------   ------------
Phi                               .17570                              .00086
Cramer's V                        .17570                              .00086
```

**Chi-Square Analysis**: **gender** (2 levels) by the helper's occupation (**occupat**, 5 levels). Deleted from this sample is the DTS level of **occupat**, thus fewer than 537 subjects.

GENDER by OCCUPAT HELPER OCCUPATION

OCCUPAT — Page 1 of 1

| GENDER | Count / Exp Val | PROFESSIONAL 1 | SERVICE/SUPPORT 2 | BLUE COLLAR 3 | UNEMPLOYED/RETIR 4 | STUDENT 5 | Row Total |
|---|---|---|---|---|---|---|---|
| FEMALE | 1 | 71 | 80 | 5 | 29 | 129 | 314 |
| | | 81.0 | 72.4 | 11.7 | 26.4 | 122.7 | 61.3% |
| MALE | 2 | 61 | 38 | 14 | 14 | 71 | 198 |
| | | 51.0 | 45.6 | 7.3 | 16.6 | 77.3 | 38.7% |
| | Column Total | 132 | 118 | 19 | 43 | 200 | 512 |
| | | 25.8% | 23.0% | 3.7% | 8.4% | 39.1% | 100.0% |

| Chi-Square | Value | DF | Significance |
|---|---|---|---|
| Pearson | 16.59292 | 4 | .00232 |
| Likelihood Ratio | 16.37661 | 4 | .00255 |
| Mantel-Haenszel test for linear association | 2.16185 | 1 | .14147 |

Minimum Expected Frequency - 7.348

| Statistic | Value | ASE1 | Val/ASE0 | Approximate Significance |
|---|---|---|---|---|
| Phi | .18002 | | | .00232 |
| Cramer's V | .18002 | | | .00232 |

**Chi-Square Analysis**: **income** (5 levels) by the helper's ethnicity (**ethnic**, 5 levels). With 25 cells, interpretation requires some careful observation. Preliminary evidence suggests that Caucasians are over-represented in the highest income categories, Hispanics and Blacks tend to be underrepresented and little difference between observed and expected values for Asians.

ETHNIC by INCOME

INCOME — Page 1 of 1

| ETHNIC | Count / Exp Val | <15,000 (1) | <25,000 (2) | <50,000 (3) | >50,000 (4) | DTS (5) | Row Total |
|---|---|---|---|---|---|---|---|
| 1 CAUCASIAN | Count | 35 | 25 | 48 | 99 | 67 | 274 |
| | Exp Val | 39.1 | 27.3 | 55.7 | 84.0 | 68.0 | 53.5% |
| 2 BLACK | Count | 3 | 10 | 11 | 13 | 13 | 50 |
| | Exp Val | 7.1 | 5.0 | 10.2 | 15.3 | 12.4 | 9.8% |
| 3 HISPANIC | Count | 11 | 10 | 23 | 17 | 17 | 78 |
| | Exp Val | 11.1 | 7.8 | 15.8 | 23.9 | 19.3 | 15.2% |
| 4 ASIAN | Count | 18 | 4 | 9 | 20 | 19 | 70 |
| | Exp Val | 10.0 | 7.0 | 14.2 | 21.5 | 17.4 | 13.7% |
| 5 OTHER/DTS | Count | 6 | 2 | 13 | 8 | 11 | 40 |
| | Exp Val | 5.7 | 4.0 | 8.1 | 12.3 | 9.9 | 7.8% |
| Column Total | | 73 | 51 | 104 | 157 | 127 | 512 |
| | | 14.3% | 10.0% | 20.3% | 30.7% | 24.8% | 100.0% |

| Chi-Square | Value | DF | Significance |
|---|---|---|---|
| Pearson | 33.82770 | 16 | .00573 |
| Likelihood Ratio | 32.31631 | 16 | .00909 |
| Mantel-Haenszel test for linear association | 2.15573 | 1 | .14204 |

Minimum Expected Frequency - 3.984

Cells with Expected Frequency < 5 - 2 OF 25 ( 8.0%)

| Statistic | Value | ASE1 | Val/ASE0 | Approximate Significance |
|---|---|---|---|---|
| Phi | .25704 | | | .00573 |
| Cramer's V | .12852 | | | .00573 |

**Chi-Square Analysis**: Tests whether those who scored below the mean on amount of help given (the not helpful group) and those who scored above the mean (the helpful group) (**cathelp**) is likely to be associated with their level of **income**. The Chi-square value suggests barely significant results (p = .05). Analysis of observed versus expected values suggests that the unhelpful group is less likely to have a high income and less likely to actually state their income. The helpful group appears to be over-represented in the highest income group and more likely to state their income.

CATHELP by INCOME

INCOME — Page 1 of 1

| CATHELP | Count / Exp Val | <15,000 1 | <25,000 2 | <50,000 3 | >50,000 4 | DTS 5 | Row Total |
|---|---|---|---|---|---|---|---|
| 1.00 NOT HELPFUL | Count | 31 | 28 | 53 | 65 | 73 | 250 |
| | Exp Val | 35.6 | 24.9 | 50.8 | 76.7 | 62.0 | 48.8% |
| 2.00 HELPFUL | Count | 42 | 23 | 51 | 92 | 54 | 262 |
| | Exp Val | 37.4 | 26.1 | 53.2 | 80.3 | 65.0 | 51.2% |
| Column Total | | 73 | 51 | 104 | 157 | 127 | 512 |
| | | 14.3% | 10.0% | 20.3% | 30.7% | 24.8% | 100.0% |

| Chi-Square | Value | DF | Significance |
|---|---|---|---|
| Pearson | 9.39593 | 4 | .05193 |
| Likelihood Ratio | 9.43173 | 4 | .05117 |
| Mantel-Haenszel test for linear association | 1.18501 | 1 | .27634 |

| | |
|---|---|
| Minimum Expected Frequency - | 24.902 |

| Statistic | Value | ASE1 | Val/ASE0 | Approximate Significance |
|---|---|---|---|---|
| Phi | .13547 | | | .05193 *1 |
| Cramer's V | .13547 | | | .05193 *1 |

**Independent-samples t-test**: Are there **gender** differences on the closeness of the relationship with the friend (**hclose**) they helped [coded on a distant(1) to close(7) scale]?

```
Number
 Variable                   of Cases       Mean          SD   SE of Mean
 -----------------------------------------------------------------------
 HCLOSE  CLOSENESS OF THE RELATIONSHIP
 FEMALE                       325        5.4769       1.494         .083
 MALE                         212        4.9670       1.491         .102
 -----------------------------------------------------------------------
          Mean Difference = .5099
          Levene's Test for Equality of Variances: F= 1.623  P= .203

       t-test for Equality of Means                                      95%
 Variances   t-value      df    2-Tail Sig     SE of Diff           CI for Diff
 -------------------------------------------------------------------------------
 Equal          3.87      535          .000           .132          (.251, .769)
 Unequal        3.87   451.79          .000           .132          (.251, .769)
 -------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on ratings of problem severity (**hseveret**) [coded on a mild(1) to severe(7) scale]?

```
Number
 Variable                   of Cases       Mean          SD   SE of Mean
 -----------------------------------------------------------------------
 HSEVERET  HELPER SEVERITY RATING
 FEMALE                       325        5.1898       1.501         .083
 MALE                         212        4.7476       1.662         .114
 -----------------------------------------------------------------------
          Mean Difference = .4422
          Levene's Test for Equality of Variances: F= 2.346  P= .126

       t-test for Equality of Means                                      95%
 Variances   t-value      df    2-Tail Sig     SE of Diff           CI for Diff
 -------------------------------------------------------------------------------
 Equal          3.20      535          .001           .138          (.171, .714)
 Unequal        3.13   418.23          .002           .141          (.164, .720)
 -------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on ratings of the controllability of the problem cause (**hcontrot**) [coded uncontrollable(1) to controllable(7)]?

```
                             Number
 Variable                   of Cases       Mean          SD   SE of Mean
 -----------------------------------------------------------------------
 HCONTROT  HELPER RATING OF CONTROLLABILITY/FAULT
 FEMALE                       325        2.9969       1.690         .094
 MALE                         212        3.5401       1.744         .120
 -----------------------------------------------------------------------
          Mean Difference = -.5432
          Levene's Test for Equality of Variances: F= .035   P= .852

       t-test for Equality of Means                                      95%
 Variances   t-value      df    2-Tail Sig     SE of Diff           CI for Diff
 -------------------------------------------------------------------------------
 Equal         -3.60      535          .000           .151        (-.840, -.246)
 Unequal       -3.57   440.94          .000           .152        (-.842, -.244)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on the amount of anger felt toward their needy friend (**angert**) [coded little(1) to much(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
ANGERT  MEAN RATING OF FOUR ANGER QUESTIONS
FEMALE                       325        2.0369       1.539         .085
MALE                         212        2.4217       1.515         .104
-----------------------------------------------------------------------
         Mean Difference = -.3848
         Levene's Test for Equality of Variances: F= .642   P= .423

      t-test for Equality of Means                                      95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -2.85      535          .005           .135        (-.650, -.119)
Unequal       -2.86   456.11          .004           .135        (-.649, -.120)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on the helpers' rating of how well the recipient is coping (**hcopet**) [coded coping poorly(1) to coping well(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
HCOPET  MEAN RATING OF 3 HELPER COPING QUESTIONS
FEMALE                       325        5.0788       1.215         .067
MALE                         212        4.7849       1.205         .083
-----------------------------------------------------------------------
         Mean Difference = .2939
         Levene's Test for Equality of Variances: F= .015   P= .903

      t-test for Equality of Means                                      95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.75      535          .006           .107          (.084, .504)
Unequal        2.75   453.78          .006           .107          (.084, .504)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on the amount of efficacy felt in the helping context (**effict**) [coded low efficacy(1) to high efficacy(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
EFFICT  MEAN OF 14 EFFICACY MEASURES
FEMALE                       325        4.8136        .966         .054
MALE                         212        4.5664        .943         .065
-----------------------------------------------------------------------
         Mean Difference = .2473
         Levene's Test for Equality of Variances: F= .718   P= .397

      t-test for Equality of Means                                      95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal          2.93      535          .004           .084          (.081, .413)
Unequal        2.94   458.64          .003           .084          (.082, .412)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **gender** differences on self ratings of empathic tendency (**empathyt**) [coded low empathy(1) to high empathy(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| EMPATHYT MEAN OF 14 EMPATHY QUESTIONS | | | | |
| FEMALE | 325 | 5.2126 | .905 | .050 |
| MALE | 212 | 4.6080 | .871 | .060 |

Mean Difference = .6047
Levene's Test for Equality of Variances: F= .159 P= .690

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | 7.68 | 535 | .000 | .079 | (.450, .759) |
| Unequal | 7.74 | 463.10 | .000 | .078 | (.451, .758) |

**Independent-samples t-test**: Are there **gender** differences on the amount of sympathy experienced in the helping context (**sympathi**) [coded low sympathy(1) to high(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| SYMPATHI SYMPATHY MEASURE | | | | |
| FEMALE | 325 | 5.3651 | 1.242 | .069 |
| MALE | 212 | 4.7909 | 1.321 | .091 |

Mean Difference = .5742
Levene's Test for Equality of Variances: F= 1.764 P= .185

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | 5.11 | 535 | .000 | .112 | (.353, .795) |
| Unequal | 5.04 | 431.32 | .000 | .114 | (.350, .798) |

**Independent-samples t-test**: Are there **cathelp** differences on the rating of the closeness of the friendship (**hclose**) [coded distant(1) to close(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| HCLOSE HELPER RATES CLOSENESS OF RELATIONSHIP | | | | |
| NOT HELPFUL | 265 | 4.9094 | 1.554 | .095 |
| HELPFUL | 272 | 5.6324 | 1.382 | .084 |

Mean Difference = -.7229
Levene's Test for Equality of Variances: F= 1.979 P= .160

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | -5.70 | 535 | .000 | .127 | (-.972, -.474) |
| Unequal | -5.69 | 524.24 | .000 | .127 | (-.972, -.473) |

**Independent-samples t-test**: Are there **cathelp** differences on the rating of problem severity (**hseveret**) [coded mild(1) to severe(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
HSEVERET  HELPER MEAN SEVERITY RATING
NOT HELPFUL                  265        4.5094       1.635         .100
HELPFUL                      272        5.5081       1.356         .082
-----------------------------------------------------------------------
         Mean Difference = -.9987
         Levene's Test for Equality of Variances: F= 11.676 P= .001

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -7.71      535          .000           .129       (-1.253, -.744)
Unequal       -7.69   512.30          .000           .130       (-1.254, -.744)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **cathelp** differences on the amount of worry experienced concerning their friend's problem (**worry**) [coded little(1) to much(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
WORRY  HELPER RATING OF WORRY EXPERIENCED BY HE
NOT HELPFUL                  265        3.6453       1.831         .112
HELPFUL                      272        4.3493       1.874         .114
-----------------------------------------------------------------------
         Mean Difference = -.7040
         Levene's Test for Equality of Variances: F= .001   P= .979

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -4.40      535          .000           .160       (-1.018, -.390)
Unequal       -4.40   535.00          .000           .160       (-1.018, -.390)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **cathelp** differences on the amount of obligation felt concerning their friend's problem (**obligat**) [coded little(1) to much(7)]?

```
                            Number
Variable                   of Cases       Mean          SD   SE of Mean
-----------------------------------------------------------------------
OBLIGAT  HELPER RATING OF OBLIGATION FELT
NOT HELPFUL                  265        4.1585       1.768         .109
HELPFUL                      272        4.9816       1.885         .114
-----------------------------------------------------------------------
         Mean Difference = -.8231
         Levene's Test for Equality of Variances: F= .827   P= .364

      t-test for Equality of Means                                       95%
Variances   t-value       df    2-Tail Sig     SE of Diff           CI for Diff
-------------------------------------------------------------------------------
Equal         -5.22      535          .000           .158       (-1.133, -.513)
Unequal       -5.22   534.24          .000           .158       (-1.133, -.513)
-------------------------------------------------------------------------------
```

**Independent-samples t-test**: Are there **cathelp** differences on the efficacy of the person trying to help (**effict**) [coded low efficacy(1) to high efficacy(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| EFFICT MEAN OF 14 EFFICACY MEASURES | | | | |
| NOT HELPFUL | 265 | 4.2876 | .909 | .056 |
| HELPFUL | 272 | 5.1334 | .822 | .050 |

Mean Difference = -.8458
Levene's Test for Equality of Variances: F= 3.708 P= .055

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | -11.31 | 535 | .000 | .075 | (-.993, -.699) |
| Unequal | -11.30 | 526.60 | .000 | .075 | (-.993, -.699) |

**Independent-samples t-test**: Are there **cathelp** differences on a self-rating of empathic tendency (**empathyt**) [coded low empathy(1) to high empathy(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| EMPATHYT MEAN OF 14 EMPATHY QUESTIONS | | | | |
| NOT HELPFUL | 265 | 4.8075 | .958 | .059 |
| HELPFUL | 272 | 5.1360 | .892 | .054 |

Mean Difference = -.3285
Levene's Test for Equality of Variances: F= .900 P= .343

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | -4.11 | 535 | .000 | .080 | (-.485, -.172) |
| Unequal | -4.11 | 530.03 | .000 | .080 | (-.485, -.171) |

**Independent-samples t-test**: Are there **cathelp** differences on the amount of sympathy experienced concerning their friend's problem (**sympathi**) [coded little(1) to much(7)]?

| Variable | Number of Cases | Mean | SD | SE of Mean |
|---|---|---|---|---|
| SYMPATHI SYMPATHY EXPERIENCED BY HELPER | | | | |
| NOT HELPFUL | 265 | 4.7547 | 1.333 | .082 |
| HELPFUL | 272 | 5.5123 | 1.159 | .070 |

Mean Difference = -.7575
Levene's Test for Equality of Variances: F= 5.086 P= .025

t-test for Equality of Means

| Variances | t-value | df | 2-Tail Sig | SE of Diff | 95% CI for Diff |
|---|---|---|---|---|---|
| Equal | -7.03 | 535 | .000 | .108 | (-.969, -.546) |
| Unequal | -7.02 | 520.80 | .000 | .108 | (-.970, -.546) |

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the **age** of the helper?

```
                         - - - - -  O N E W A Y  - - - - -
     Variable  AGE
  By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                   Sum of        Mean            F      F
       Source           D.F.     Squares       Squares         Ratio  Prob.
Between Groups            3     6068.2652     2022.7551      10.8655  .0000
Within Groups           533    99225.0011      186.1632
Total                   536   105293.2663

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 9.6479 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R C
                          O E A I
                          A L T L
                          L A A L
                            T S N
                          D I T E
    Mean      PROBLEM
   29.4279    GOAL DIS
   30.1770    RELATION
   34.3077    CATASTRO
   38.8372    ILLNESS     * *
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the rating of problem severity (**hseveret**)?

```
                        - - - - -  O N E W A Y  - - - - -

     Variable  HSEVERET   PROBLEM SEVERITY
  By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                   Sum of        Mean            F      F
       Source           D.F.     Squares       Squares         Ratio  Prob.
Between Groups            3       99.9638       33.3213      14.3453  .0000
Within Groups           533     1238.0510        2.3228
Total                   536     1338.0148

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 1.0777 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R   C
                          O E I A
                          A L L T
                          L A L A
                            T N S
                          D I E T
    Mean      PROBLEM
    4.5817    GOAL DIS
    5.1459    RELATION    *
    5.7035    ILLNESS     * *
    6.0000    CATASTRO    *
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with how much the helper worries (**worry**)?

```
                        - - - - -  O N E W A Y  - - - - -

      Variable  WORRY      HELPER RATING OF WORRY EXPERIENCED
   By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.
Between Groups             3      116.1321       38.7107      11.5469  .0000
Within Groups            533     1786.8660        3.3525
Total                    536     1902.9981

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 1.2947 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R   C
                          O E I A
                          A L L T
                          L A L A
                            T N S
                          D I E T
     Mean      PROBLEM
     3.6332    GOAL DIS
     3.9761    RELATION
     4.8372    ILLNESS    * *
     5.3846    CATASTRO   * *
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the helper's perception of how well the recipient is coping (**hcopet**)?

```
                        - - - - -  O N E W A Y  - - - - -
      Variable  HCOPET     HOW WELL THE RECIPIENT IS COPING
   By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares         Ratio  Prob.

Between Groups             3       12.2539        4.0846       2.7771  .0407
Within Groups            533      783.9412        1.4708
Total                    536      796.1951

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .8576 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                            R C G
                          I E A O
                          L L T A
                          L A A L
                          N T S
                          E I T D
     Mean      PROBLEM
     4.6756    ILLNESS
     4.9225    RELATION
     4.9538    CATASTRO
     5.1079    GOAL DIS   *
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the amount of sympathy (sympathi) experienced by the helper?

```
                       - - - - -  O N E W A Y  - - - - -

      Variable  SYMPATHI   HELPER'S EXPERIENCE OF SYMPATHY
   By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                   Sum of         Mean           F      F
        Source            D.F.    Squares       Squares        Ratio  Prob.
Between Groups              3      52.2305      17.4102      10.8150  .0000
Within Groups             533     858.0356       1.6098
Total                     536     910.2661

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .8972 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R   C
                          O E I A
                          A L L T
                          L A L A
                            T N S
                          D I E T
     Mean      PROBLEM
     4.7962    GOAL DIS
     5.3094    RELATION   *
     5.5310    ILLNESS    *
     5.8205    CATASTRO   *
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the total amount of help given (**tothelp**)?

```
                       - - - - -  O N E W A Y  - - - - -

      Variable  TOTHELP    COMBINED HELP MEASURE--QUANTITY & QUALIT
   By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                   Sum of         Mean           F      F
        Source            D.F.    Squares       Squares        Ratio  Prob.
Between Groups              3       6.9146       2.3049       4.3395  .0049
Within Groups             533     283.0938        .5311
Total                     536     290.0084

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .5153 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R   C
                          O E I A
                          A L L T
                          L A L A
                            T N S
                          D I E T
     Mean      PROBLEM
     -.1245    GOAL DIS
      .0715    RELATION   *
      .1254    ILLNESS    *
      .2697    CATASTRO
```

**One-way ANOVA**: Is the type of problem experienced (**problem,** 4 levels) associated with the amount of time spent helping (**thelplnz**)?

```
                        - - - - -  O N E W A Y  - - - - -

      Variable  THELPLNZ   LENGTH OF TIME HELPING
   By Variable  PROBLEM    TYPE OF PROBLEM EXPERIENCED

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares        Ratio  Prob.
Between Groups             3      30.8078       10.2693      12.3555  .0000
Within Groups            533     443.0041         .8312
Total                    536     473.8119

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .6447 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          G R C
                          O E A I
                          A L T L
                          L A A L
                            T S N
                          D I T E
     Mean      PROBLEM
     -.2738    GOAL DIS
      .1729    RELATION   *
      .2672    CATASTRO   *
      .2742    ILLNESS    *
```

**One-way ANOVA**: Is the amount of education (**school,** 5 levels) associated with the rating of problem severity (**hseveret**)?

```
                     - - - - -  O N E W A Y  - - - - -

      Variable  HSEVERET   HELPER MEAN SEVERITY RATING
   By Variable  SCHOOL     NUMBER OF YEARS IN SCHOOL

                                  Sum of        Mean            F      F
        Source           D.F.    Squares       Squares        Ratio  Prob.
Between Groups             5      31.9487        6.3897       2.6315  .0231
Within Groups            528    1282.0653        2.4282
Total                    533    1314.0141

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= 1.1019 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                          1 1 1 1
                          9 5 3 7
                            - - -
                          O 1 1 1 1
                          R 6 4 8 2
     Mean      SCHOOL
     4.8561    19 OR MO
     4.8610    15-16 YR
     5.0192    13-14 YR
     5.1364    17-18 YR
     5.6915    12 YR      * * * *
```

**One-way ANOVA**: Is the amount of education (**school,** 6 levels) associated with the quality of the help given (**tqualitz**)?

```
                        - - - - -  O N E W A Y  - - - - -

      Variable  TQUALITZ   HELP QUALITY
   By Variable  SCHOOL     NUMBER OF YEARS IN SCHOOL

                                   Sum of         Mean             F      F
        Source           D.F.    Squares       Squares          Ratio   Prob.
Between Groups             5      12.4550        2.4910        3.3371  .0056
Within Groups            528     394.1326         .7465
Total                    533     406.5875

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .6109 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle
                            1 1   1 1
                          9 5 7   3 9
                          - - -   -
                          1 1 1 1 1 O
                          1 6 8 2 4 R
     Mean      SCHOOL
     -.8549    9-11 YR
     -.0941    15-16 YR   *
     -.0044    17-18 YR   *
      .0107    12 YR      *
      .1067    13-14 YR   *
      .2581    19 OR MO   * *
```

**One-way ANOVA**: Is the ethnicity of the helper (**ethnic,** 4 levels) associated with the quality of help given (**tqualitz**)?

```
- - - - -  O N E W A Y  - - - - -

      Variable  TQUALITZ   HELP QUALITY
   By Variable  ETHNIC

                                   Sum of         Mean             F      F
        Source           D.F.    Squares       Squares          Ratio   Prob.
Between Groups             3       7.3158        2.4386        3.1267  .0255
Within Groups            489     381.3880         .7799
Total                    492     388.7038

Multiple Range Tests:  LSD test with significance level .05
The difference between two means is significant if
  MEAN(J)-MEAN(I)  >= .6245 * RANGE * SQRT(1/N(I) + 1/N(J))
  with the following value(s) for RANGE: 2.78
   (*) Indicates significant differences which are shown in the lower triangle

                            C H
                            A I
                            U S
                          A C P B
                          S A A L
                          I S N A
     Mean      ETHNIC
     -.2662    ASIAN
     -.0044    CAUCASIA   *
      .0716    HISPANIC   *
      .1988    BLACK      *
```

**Three-way ANOVA with 2 ´ main effects and no interactions**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on the length of time spent helping (**thelplnz**)?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 THELPLNZ LENGTH OF TIME HELPING
            by   GENDER
                 ETHNIC
                 PROBLEM  TYPE OF PROBLEM EXPERIENCED

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 46.951 | 7 | 6.707 | 8.037 | .000 |
| **GENDER** | **16.661** | **1** | **16.661** | **19.964** | **.000** |
| ETHNIC | 1.446 | 3 | .482 | .578 | .630 |
| **PROBLEM** | **28.844** | **3** | **9.615** | **11.521** | **.000** |
| 2-Way Interactions | 2.706 | 15 | .180 | .216 | .999 |
| GENDER ETHNIC | .866 | 3 | .289 | .346 | .792 |
| GENDER PROBLEM | .125 | 3 | .042 | .050 | .985 |
| ETHNIC PROBLEM | 1.822 | 9 | .202 | .243 | .988 |
| 3-Way Interactions | 7.512 | 8 | .939 | 1.125 | .345 |
| GENDER ETHNIC PROBLEM | 7.512 | 8 | .939 | 1.125 | .345 |
| Explained | 57.169 | 30 | 1.906 | 2.283 | .000 |
| Residual | 385.552 | 462 | .835 | | |
| Total | 442.721 | 492 | .900 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with 3 ´ main effects and no interactions**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on the total amount of help given (**tothelp**)?

* * * A N A L Y S I S  O F  V A R I A N C E * * *

TOTHELP COMBINED HELP MEASURE--QUANTITY & QUALIT
by GENDER
ETHNIC
PROBLEM TYPE OF PROBLEM EXPERIENCED

HIERARCHICAL sums of squares

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 20.671 | 7 | 2.953 | 5.591 | .000 |
| **GENDER** | **10.372** | **1** | **10.372** | **19.639** | **.000** |
| **ETHNIC** | **3.683** | **3** | **1.228** | **2.324** | **.074** |
| **PROBLEM** | **6.615** | **3** | **2.205** | **4.175** | **.006** |
| 2-Way Interactions | 4.077 | 15 | .272 | .515 | .933 |
| GENDER ETHNIC | .719 | 3 | .240 | .454 | .715 |
| GENDER PROBLEM | .900 | 3 | .300 | .568 | .636 |
| ETHNIC PROBLEM | 2.113 | 9 | .235 | .445 | .910 |
| 3-Way Interactions | 4.165 | 8 | .521 | .986 | .446 |
| GENDER ETHNIC PROBLEM | 4.165 | 8 | .521 | .986 | .446 |
| Explained | 28.912 | 30 | .964 | 1.825 | .006 |
| Residual | 244.001 | 462 | .528 | | |
| Total | 272.913 | 492 | .555 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with 3 ´ main effects and no interactions**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on the rating of problem severity (**hseveret**)?

```
            * * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 HSEVERET RATING OF PROBLEM SEVERITY
            by   GENDER
                 ETHNIC
                 PROBLEM  TYPE OF PROBLEM EXPERIENCED

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 116.037 | 7 | 16.577 | 7.566 | .000 |
| GENDER | 26.571 | 1 | 26.571 | 12.128 | .001 |
| ETHNIC | 4.532 | 3 | 1.511 | .689 | .559 |
| PROBLEM | 84.935 | 3 | 28.312 | 12.923 | .000 |
| 2-Way Interactions | 21.670 | 15 | 1.445 | .659 | .824 |
| GENDER ETHNIC | 3.898 | 3 | 1.299 | .593 | .620 |
| GENDER PROBLEM | 1.086 | 3 | .362 | .165 | .920 |
| ETHNIC PROBLEM | 15.410 | 9 | 1.712 | .782 | .634 |
| 3-Way Interactions | 33.448 | 8 | 4.181 | 1.908 | .057 |
| GENDER ETHNIC PROBLEM | 33.448 | 8 | 4.181 | 1.908 | .057 |
| Explained | 171.156 | 30 | 5.705 | 2.604 | .000 |
| Residual | 1012.156 | 462 | 2.191 | | |
| Total | 1183.312 | 492 | 2.405 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with 1 ´ main effect and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on the amount of anger felt toward the friend (**angert**)?

```
            * * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 ANGERT    MEAN RATING OF FOUR ANGER QUESTIONS
            by   GENDER
                 ETHNIC
                 PROBLEM   TYPE OF PROBLEM EXPERIENCED

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 38.718 | 7 | 5.531 | 2.484 | .016 |
| **GENDER** | **20.704** | **1** | **20.704** | **9.297** | **.002** |
| ETHNIC | 3.977 | 3 | 1.326 | .595 | .618 |
| PROBLEM | 14.037 | 3 | 4.679 | 2.101 | .099 |
| 2-Way Interactions | 64.439 | 15 | 4.296 | 1.929 | .019 |
| GENDER ETHNIC | 11.563 | 3 | 3.854 | 1.731 | .160 |
| GENDER PROBLEM | 15.208 | 3 | 5.069 | 2.276 | .079 |
| **ETHNIC PROBLEM** | **48.205** | **9** | **5.356** | **2.405** | **.011** |
| 3-Way Interactions | 23.632 | 8 | 2.954 | 1.326 | .228 |
| GENDER ETHNIC PROBLEM | 23.632 | 8 | 2.954 | 1.326 | .228 |
| Explained | 126.788 | 30 | 4.226 | 1.898 | .003 |
| Residual | 1028.855 | 462 | 2.227 | | |
| Total | 1155.643 | 492 | 2.349 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with no main effects and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on a feeling of obligation toward the friend (**obligat**)?

```
            * * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 OBLIGAT  HELPER RATING OF OBLIGATION FELT
            by   GENDER
                 ETHNIC
                 PROBLEM  TYPE OF PROBLEM EXPERIENCED

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 11.984 | 7 | 1.712 | .497 | .837 |
| GENDER | .020 | 1 | .020 | .006 | .939 |
| ETHNIC | 5.416 | 3 | 1.805 | .524 | .666 |
| PROBLEM | 6.548 | 3 | 2.183 | .633 | .594 |
| 2-Way Interactions | 80.342 | 15 | 5.356 | 1.554 | .083 |
| GENDER ETHNIC | 17.777 | 3 | 5.926 | 1.719 | .162 |
| GENDER PROBLEM | 2.328 | 3 | .776 | .225 | .879 |
| **ETHNIC PROBLEM** | **58.486** | **9** | **6.498** | **1.885** | **.052** |
| 3-Way Interactions | 10.577 | 8 | 1.322 | .384 | .929 |
| GENDER ETHNIC PROBLEM | 10.577 | 8 | 1.322 | .384 | .929 |
| Explained | 102.903 | 30 | 3.430 | .995 | .476 |
| Residual | 1592.330 | 462 | 3.447 | | |
| Total | 1695.233 | 492 | 3.446 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with 1 ´ main effect and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and problem type (**problem**, 4 levels) have on a self-rating of empathic tendency (**empathyt**)?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

     EMPATHYT MEAN OF 14 EMPATHY QUESTIONS
by   GENDER
     ETHNIC
     PROBLEM  TYPE OF PROBLEM EXPERIENCED

     HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 45.845 | 7 | 6.549 | 8.617 | .000 |
| **GENDER** | **42.277** | **1** | **42.277** | **55.624** | **.000** |
| ETHNIC | 3.016 | 3 | 1.005 | 1.323 | .266 |
| PROBLEM | .552 | 3 | .184 | .242 | .867 |
| 2-Way Interactions | 14.525 | 15 | .968 | 1.274 | .214 |
| **GENDER ETHNIC** | **7.086** | **3** | **2.362** | **3.108** | **.026** |
| GENDER PROBLEM | .236 | 3 | .079 | .103 | .958 |
| ETHNIC PROBLEM | 8.269 | 9 | .919 | 1.209 | .287 |
| 3-Way Interactions | 5.361 | 8 | .670 | .882 | .532 |
| GENDER ETHNIC PROBLEM | 5.361 | 8 | .670 | .882 | .532 |
| Explained | 65.732 | 30 | 2.191 | 2.883 | .000 |
| Residual | 351.145 | 462 | .760 | | |
| Total | 416.877 | 492 | .847 | | |

537 cases were processed.
44 cases (8.2 pct) were missing.

**Three-way ANOVA with 3 ´ main effects and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and **income**, (5 levels) have on a the amount of time spent helping (**thelplnz**)?

* * * A N A L Y S I S O F V A R I A N C E * * *

THELPLNZ TOTAL TIME HELPING
by GENDER
PROBLEM TYPE OF PROBLEM EXPERIENCED
INCOME

HIERARCHICAL sums of squares

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 51.828 | 8 | 6.479 | 8.127 | .000 |
| **GENDER** | **15.719** | **1** | **15.719** | **19.720** | **.000** |
| **PROBLEM** | **25.712** | **3** | **8.571** | **10.752** | **.000** |
| **INCOME** | **10.397** | **4** | **2.599** | **3.261** | **.012** |
| 2-Way Interactions | 18.102 | 19 | .953 | 1.195 | .256 |
| GENDER PROBLEM | .867 | 3 | .289 | .362 | .780 |
| **GENDER INCOME** | **10.233** | **4** | **2.558** | **3.209** | **.013** |
| PROBLEM INCOME | 8.553 | 12 | .713 | .894 | .553 |
| 3-Way Interactions | 5.315 | 9 | .591 | .741 | .671 |
| GENDER PROBLEM INCOME | 5.315 | 9 | .591 | .741 | .671 |
| Explained | 75.245 | 36 | 2.090 | 2.622 | .000 |
| Residual | 398.567 | 500 | .797 | | |
| Total | 473.812 | 536 | .884 | | |

537 cases were processed.
0 cases (.0 pct) were missing.

**Three-way ANOVA with 1 ´ main effects and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and **income**, (5 levels) have on the quality of help given (**tqualitz**)?

```
            * * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 TQUALITZ QUALITY OF HELP
            by   GENDER
                 PROBLEM  TYPE OF PROBLEM EXPERIENCED
                 INCOME

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 13.670 | 8 | 1.709 | 2.244 | .023 |
| **GENDER** | **7.326** | **1** | **7.326** | **9.621** | **.002** |
| PROBLEM | 2.247 | 3 | .749 | .984 | .400 |
| INCOME | 4.097 | 4 | 1.024 | 1.345 | .252 |
| 2-Way Interactions | 13.619 | 19 | .717 | .941 | .531 |
| GENDER PROBLEM | 2.319 | 3 | .773 | 1.015 | .386 |
| **GENDER INCOME** | **7.746** | **4** | **1.936** | **2.543** | **.039** |
| PROBLEM INCOME | 4.036 | 12 | .336 | .442 | .946 |
| 3-Way Interactions | 5.079 | 9 | .564 | .741 | .671 |
| GENDER PROBLEM INCOME | 5.079 | 9 | .564 | .741 | .671 |
| Explained | 32.367 | 36 | .899 | 1.181 | .222 |
| Residual | 380.751 | 500 | .762 | | |
| Total | 413.118 | 536 | .771 | | |

```
537 cases were processed.
0 cases (.0 pct) were missing.
```

**Three-way ANOVA with 3 ´ main effects and 1 ´ 2-way interaction**: What influence does **gender** (2 levels), **ethnic** (4 levels), and **income**, (5 levels) have on the amount of total help given (**tothelp**)?

```
            * * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 TOTHELP  COMBINED HELP MEASURE--QUANTITY & QUALIT
            by   GENDER
                 PROBLEM  TYPE OF PROBLEM EXPERIENCED
                 INCOME

                 HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 22.864 | 8 | 2.858 | 5.549 | .000 |
| **GENDER** | **11.127** | **1** | **11.127** | **21.606** | **.000** |
| **PROBLEM** | **5.179** | **3** | **1.726** | **3.352** | **.019** |
| **INCOME** | **6.558** | **4** | **1.639** | **3.183** | **.013** |
| 2-Way Interactions | 7.237 | 19 | .381 | .740 | .778 |
| GENDER PROBLEM | .558 | 3 | .186 | .361 | .781 |
| **GENDER INCOME** | **5.148** | **4** | **1.287** | **2.499** | **.042** |
| PROBLEM INCOME | 2.120 | 12 | .177 | .343 | .981 |
| 3-Way Interactions | 2.406 | 9 | .267 | .519 | .861 |
| GENDER PROBLEM INCOME | 2.406 | 9 | .267 | .519 | .861 |
| Explained | 32.507 | 36 | .903 | 1.753 | .005 |
| Residual | 257.501 | 500 | .515 | | |
| Total | 290.008 | 536 | .541 | | |

```
537 cases were processed.
0 cases (.0 pct) were missing.
```

**Three-way ANOVA with 2 ´ main effects and 2 ´ 2-way interactions**: What influence does **gender** (2 levels), **problem** type (4 levels), and **income**, (5 levels) have on the amount of anger experienced by the helper (**angert**)?

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

       ANGERT   MEAN RATING OF FOUR ANGER QUESTIONS
   by  GENDER
       PROBLEM  TYPE OF PROBLEM EXPERIENCED
       INCOME

       HIERARCHICAL sums of squares
```

| Source of Variation | Sum of Squares | DF | Mean Square | F | Sig of F |
|---|---|---|---|---|---|
| Main Effects | 49.865 | 8 | 6.233 | 2.808 | .005 |
| **GENDER** | **18.996** | **1** | **18.996** | **8.556** | **.004** |
| PROBLEM | 8.769 | 3 | 2.923 | 1.317 | .268 |
| **INCOME** | **22.100** | **4** | **5.525** | **2.489** | **.043** |
| 2-Way Interactions | 83.385 | 19 | 4.389 | 1.977 | .008 |
| GENDER PROBLEM | 4.347 | 3 | 1.449 | .653 | .582 |
| **GENDER INCOME** | **36.362** | **4** | **9.090** | **4.095** | **.003** |
| **PROBLEM INCOME** | **46.367** | **12** | **3.864** | **1.740** | **.056** |
| 3-Way Interactions | 28.038 | 9 | 3.115 | 1.403 | .184 |
| GENDER PROBLEM INCOME | 28.038 | 9 | 3.115 | 1.403 | .184 |
| Explained | 161.288 | 36 | 4.480 | 2.018 | .001 |
| Residual | 1110.025 | 500 | 2.220 | | |
| Total | 1271.313 | 536 | 2.372 | | |

537 cases were processed.
0 cases (.0 pct) were missing.

**Multiple Regression Analysis**: Criterion variable: Total help (**tothelp**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

TOTAL HELP FOR ALL

Equation Number 1    Dependent Variable..   TOTHELP   COMBINED HELP MEASURE--QUA

Variable(s) Entered on Step Number
   6..    OBLIGAT   HELPER RATING OF OBLIGATION FELT

Multiple R           .65595
R Square             .43027
Adjusted R Square    .42382
Standard Error       .55834

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           6           124.78215         20.79702
Residual           530           165.22629           .31175

F =      66.71107       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta          T  Sig T
HCLOSE          .049710     .017191    .102175      2.892  .0040
HSEVERET        .112620     .015923    .241902      7.073  .0000
OBLIGAT         .031165     .013578    .079333      2.295  .0221
HCOPET          .051504     .020607    .085339      2.499  .0127
EFFICT          .346784     .026728    .454256     12.974  .0000
EMPATHYT        .076205     .026452    .097258      2.881  .0041
(Constant)    -3.238400     .190231               -17.024  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
GENDER      -.057742 -.070785    .851015    -1.632  .1032
AGE          .056900  .070799    .815346     1.632  .1032
INCOME      -.058216 -.076707    .859669    -1.769  .0774
HCONTROT    -.054860 -.070696    .860978    -1.630  .1037
ANGERT      -.029735 -.036854    .836376     -.848  .3967
WORRY        .042053  .047165    .716661     1.086  .2780
SYMPATHI     .050358  .057078    .731931     1.315  .1891


End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: For men, criterion variable: Total help (**tothelp**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

TOTAL HELP FOR MEN

Equation Number 1    Dependent Variable..   TOTHELP   COMBINED HELP MEASURE--QUA

Variable(s) Entered on Step Number
   4..    EMPATHYT  MEAN OF 14 EMPATHY QUESTIONS

Multiple R           .63536
R Square             .40368
Adjusted R Square    .39216
Standard Error       .58462

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           4            47.89344         11.97336
Residual           207            70.74820           .34178
F =      35.03249       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
HSEVERET        .113502     .024633    .251551     4.608  .0000
HCONTROT       -.063299     .023598   -.147180    -2.682  .0079
EFFICT          .393926     .043812    .495392     8.991  .0000
EMPATHYT        .096196     .047694    .111766     2.017  .0450
(Constant)    -2.733743     .303802               -8.998  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
AGE          .080921  .102847    .935531     1.484  .1393
INCOME       .036027  .045544    .907399      .654  .5136
HCLOSE       .004726  .005824    .905628      .084  .9335
ANGERT      -.010224 -.011330    .732313     -.163  .8710
WORRY        .070692  .076172    .692353     1.096  .2742
OBLIGAT      .039267  .047425    .869847      .681  .4964
HCOPET       .099731  .126775    .936320     1.834  .0680
SYMPATHI     .053668  .055915    .647299      .804  .4224


End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: For women, criterion variable: Total help (**tothelp**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *
```

**TOTAL HELP FOR WOMEN**

```
Variable(s) Entered on Step Number
   5..    OBLIGAT   HELPER RATING OF OBLIGATION FELT


Multiple R           .67394
R Square             .45420
Adjusted R Square    .44564
Standard Error       .52361

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           5            72.78014         14.55603
Residual           319            87.45952           .27417
F =      53.09168       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
INCOME         -.069339     .021967   -.131726    -3.157  .0017
HCLOSE          .086755     .020601    .184289     4.211  .0000
HSEVERET        .097910     .019835    .209005     4.936  .0000
OBLIGAT         .047716     .015605    .131467     3.058  .0024
EFFICT          .345311     .031439    .474165    10.984  .0000
(Constant)    -2.498605     .202708              -12.326  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
AGE          .010194  .012780    .847936      .228  .8199
HCONTROT     .003878  .005242    .892479      .093  .9256
ANGERT     -.024643 -.032672    .875072     -.583  .5603
WORRY        .023752  .028274    .773421      .504  .6143
HCOPET       .046513  .059548    .869952     1.064  .2882
EMPATHYT     .024154  .031701    .888473      .566  .5721
SYMPATHI     .006835  .008236    .792561      .147  .8833


End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: Criterion variable: Help quality (**tqualitz**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

HELP QUALITY FOR ALL

Equation Number 1    Dependent Variable..   TQUALITZ   MEAN OF HELPER/RECIPIENT


Variable(s) Entered on Step Number
   5..    OBLIGAT   HELPER RATING OF OBLIGATION FELT


Multiple R           .59006
R Square             .34817
Adjusted R Square    .34204
Standard Error       .71212

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           5           143.83710         28.76742
Residual           531           269.28130           .50712
F =      56.72693       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
HSEVERET        .068772     .020182    .123768     3.408  .0007
ANGERT         -.088324     .021242   -.154941    -4.158  .0000
OBLIGAT         .041397     .016907    .088291     2.449  .0147
HCOPET          .143940     .027347    .199826     5.263  .0000
EFFICT          .385048     .033440    .422595    11.515  .0000
(Constant)    -2.869453     .219401              -13.079  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
AGE          .017948  .021614    .849026      .498  .6189
INCOME      -.028559 -.035354    .851661     -.814  .4158
HCLOSE       .036316  .041645    .840813      .960  .3377
HCONTROT    -.061403 -.065886    .695635    -1.520  .1291
WORRY       -.055659 -.058303    .715211    -1.345  .1794
EMPATHYT     .066361  .080004    .846071     1.848  .0652
SYMPATHI     .017014  .017821    .715121      .410  .6817


End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: For men, criterion variable: Help quality (**tqualitz**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *
```

**HELP QUALITY FOR MEN**

```
Equation Number 1    Dependent Variable..   TQUALITZ   MEAN OF HELPER/RECIPIENT


Variable(s) Entered on Step Number
   3..    HCONTROT  HELPER MEAN RATING OF CONTROLLABILITY


Multiple R           .52991
R Square             .28081
Adjusted R Square    .27044
Standard Error       .75087

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           3            45.78859         15.26286
Residual           208           117.27132           .56380
F =      27.07120       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
HCONTROT       -.102433     .029866   -.203158    -3.430  .0007
HCOPET          .170310     .043324    .233486     3.931  .0001
EFFICT          .380393     .055139    .408050     6.899  .0000
(Constant)    -2.332161     .334365               -6.975  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
AGE          .022318  .026043    .973461      .375  .7082
INCOME      -.006509 -.007632    .973033     -.110  .9127
HCLOSE      -.012703 -.014542    .938031     -.209  .8345
HSEVERET     .100782  .116047    .953555     1.681  .0943
ANGERT      -.007687 -.007855    .751032     -.113  .9101
WORRY        .012593  .014405    .934892      .207  .8360
OBLIGAT      .037260  .041734    .900508      .601  .5485
EMPATHYT     .094822  .108485    .941394     1.570  .1179
SYMPATHI     .067645  .071871    .811854     1.037  .3011


End Block Number   1   PIN =     .060 Limits reached.
```

**Multiple Regression Analysis**: For women, criterion variable: Help quality (**tqualitz**) with predictors of **gender**, **age**, **income**, closeness of the friend (**hclose**), problem severity (**hseveret**), feelings of obligation (**obligat**), how well the recipient is coping (**hcopet**), the self-efficacy of the helper (**effict**), the empathic tendency of the helper (**empathyt**), controllability of the problem cause (**hcontrot**), anger felt by the helper (**angert**), worry felt by the helper (**worry**), and the sympathy of the helper (**sympathi**),

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *
```

**HELP QUALITY FOR WOMEN**

```
Equation Number 1    Dependent Variable..   TQUALITZ   MEAN OF HELPER/RECIPIENT


Variable(s) Entered on Step Number
   6..    HCLOSE    HELPER RATES CLOSENESS OF RELATIONSHIP


Multiple R           .63583
R Square             .40428
Adjusted R Square    .39304
Standard Error       .67433

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           6            98.13214         16.35536
Residual           318           144.60015           .45472
F =      35.96817       Signif F =  .0000


------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T
HCLOSE          .057141     .026929    .098622     2.122  .0346
HSEVERET        .066748     .025791    .115768     2.588  .0101
ANGERT         -.102673     .026793   -.182609    -3.832  .0002
OBLIGAT         .049627     .020195    .111094     2.457  .0145
HCOPET          .095803     .035270    .134531     2.716  .0070
EFFICT          .410890     .041416    .458423     9.921  .0000
(Constant)    -3.044398     .284576              -10.698  .0000


------------- Variables not in the Equation -------------

Variable     Beta In  Partial  Min Toler         T  Sig T
AGE          .016733  .019941    .761723      .355  .7227
INCOME      -.052199 -.066203    .753444    -1.181  .2384
HCONTROT     .018269  .020587    .669760      .367  .7141
WORRY       -.058573 -.065675    .748942    -1.172  .2421
EMPATHYT     .028663  .035668    .755988      .635  .5256
SYMPATHI    -.044048 -.048796    .731068     -.870  .3851


End Block Number   1   PIN =     .060 Limits reached.
```