CHAPTER 1

Join the Allyn & Bacon First Editions Club!

To start your membership in our First Editions Club, simply read the following chapter and complete a brief questionnaire at: **www.ablongman.com/swsurvey.**

As a Thank You, we will send you a Penguin-Putnam title of your choice!

We look forward to hearing from you!

Data Analysis with Spreadsheets (with CD-ROM)

David A. Patterson
Randall E. Basham

ISBN-10: 0-205-40751-X
ISBN-13: 978-0-205-40751-4

CHAPTER 1: An Introduction to Spreadsheets

ALLYN & BACON
www.ablongman.com/socialwork
Visit www.ablongman.com/replocator to contact your local Allyn & Bacon/Longman representative.

CHAPTER

1 An Introduction to Spreadsheets

Spreadsheets—First Facts

An inescapable task of the early 21st century in social service practice is the management of information. Most social workers, counselors, and human service workers likely enter their respective professional roles with the desire to provide necessary services to a particular client population, having neither the intention nor the wish to devote their time and energy to the collection, analysis, and dissemination of information. The emerging practice reality is that data have to be collected that document to whom what specific services are delivered, for what duration, in what quantity, and resulting in what outcomes. Governmental agencies, public and private insurers, managed care companies, as well as agency administrators and program evaluators, require the capture, analysis, and reporting of social services delivery information. Further, budgetary analysis and financial reporting requirements of social service agencies require skills in the evaluation of possible budget scenarios, the preparation of viable budget reports, and the monitoring of resources and expenditures over the life of the budget. Computers and the spectrum of information technology software and hardware now available are the essential tools for each of these information management tasks. Consequently, the ability to use information technology in the gathering, examination, and presentation of information regarding the provision of social services is now a requisite skill set for social service professionals.

Spreadsheet applications, or simply "spreadsheets," are computer programs that display a matrix of rows and columns of cells into which information, in the form of numbers, text, or formulas, is entered and displayed. The size of this rows and columns matrix is generally only limited by the available memory of the computer in which the spreadsheet software resides. As a result of the ever-growing power of personal computers, manifest in processing speed, storage capacity, and memory (RAM), spreadsheets are now capable of holding and manipulating vast amounts of information. Beyond the simple storage of information, the usefulness of spreadsheets stems from the ease with which they allow users to interact with data. For instance, spreadsheets can be used in social services practice settings to (a) record and graph the changes clients make over the course of service delivery, (b) collect and analyze data on how groups and individuals within the groups change over time, (c) record and analyze agency income and expenses, and (d) generate graphs and tables reporting agency service delivery and financial status to private and governmental funding sources, oversight boards, and constituents.

Spreadsheets are a rather neglected tool for the evaluation of social service practice and practice-based research. Social service practice settings have seen a growing movement toward greater accountability over the last generation. Accountability standards have been recommended for the purposes of determining intervention effectiveness, reducing risk of harm to clients, providing cost-effectiveness information on services provided, and for tracking overall program progress at various levels in social service agencies. Spreadsheets, which are commonly found on almost all modern-day computers, are a ready instrument for

accomplishing each of these data analysis tasks. The inattention in social service practice to the utility and versatility of spreadsheets is explored below.

The Brief History of Spreadsheets

The history of spreadsheets is paradoxically both long and short. Over the course of civilization dating back to the Babylonian Empire (4,500 B.C), the recording of information in ledgers or registers has been a means to account for financial transactions, inventory objects, assess taxes, and list the births, actions, and deaths of individuals. Such ledgers were pragmatic tools for recording and tracking information, both fiscal and demographic, as well as means to preserve and convey the actions and transactions of individuals, families, communities, businesses, and nations. More recently, though prior to the advent of electronic computers, accountants and others have used worksheets composed of matrices of columns and rows to record credits and debts, to note financial transactions, and to track the fiscal well-being of individuals, enterprises, and institutions.

The trouble with paper-based ledgers or spreadsheets is that they are static. A change of a number in a column or the removal or addition of information located in a row requires the recalculation of the bottom line, the manual re-entry of new figures. It is hardly surprising that prior to the introduction of the electronic spreadsheet, two of the essential tools of any accountant were an eraser and a sharp pencil, requisite instruments of the impermanent tasks of preliminary calculations. Further, it is of little wonder that with the advent of computers, researchers and academics turned their attention to how the tasks of accounting and inventory control might be automated and made more flexible.

It is generally recognized that one of the pioneers in the development of computer-based spreadsheets was Richard Mattessich. Mattessich's groundbreaking work appeared first in a paper (Mattessich, 1961) and later in two books (Mattessich, 1964a; Mattessich, 1964b). In these publications he laid out the basic elements of computerized financial spreadsheets, including the use of matrices, formulas calculating cell values, and simulations of "what if" scenarios (Gaffikin, 1996; Legg, 1988; Murphy, 1997). The fundamental limitation of Mattessich's early work was not conceptual but instead pragmatic. His methods required mainframe computers (the only type available at the time) and complex programming, both of which had significant associated costs.

The birth of modern, personal computer-based spreadsheets occurred in 1978 when a Harvard Business School student, Dan Bricklin, balked at the complexity involved in doing a "case study" project either by hand or on a mainframe computer (Power, 2000). As an alternative he envisioned an interactive electronic spreadsheet. He initially created a "working prototype" composed of 20 rows and 5 columns in which a user could input numbers and produce calculations. He then joined with Bob Frankston, whom he knew from MIT, to improve the functioning of the original prototype by adding the ability to scroll, increase the speed of calculations, enhance the arithmetic, and optimize the program code so it would run on an affordable personal computer, one of the first Apple Computer machines. The resulting program was called VisiCalc, which was short for "visible calculator."

Bricklin (2003) points out that prior to the development of VisiCalc, there existed other row and column tabulation programs. VisiCalc, however, was unique as the first "electronic spreadsheet" having interactive features including scrolling, automatic recalculation of values in cells based on stored formulas, and ease of data input, formatting, and output. The introduction of VisiCalc had a catalytic sales effect on the blossoming personal computer industry, as it was a software product with real utility for organizations in the business and finance sectors of the economy. Further, the VisiCalc user interface became the prototype for all spreadsheet programs subsequently developed. Bricklin and Frankston formed Software Arts in January, 1979 and eventually sold over 500,000 copies of this first "electronic spreadsheet." Readers interested in seeing screenshots of the original VisiCalc program should visit Bricklin's website, www.bricklin.com/firstspreadsheetquestion.htm. A

copy of the original VisiCalc is available for downloading at www.bricklin.com/history/vcexecutable.htm.

The VisiCalc dominance of the spreadsheet market was eclipsed in the early 1980s by a new product, Lotus 1-2-3, which offered an easier to use interface, the ability to create charts, and database functions. Lotus 1-2-3 soon became a best-selling software application and the de facto standard for spreadsheets during that time period (Power, 2000). In the mid-1980s Microsoft introduced Excel, which was developed for an early version of the Apple Macintosh. At that time, the Apple Macintosh was the only personal computer with a graphical user interface and a pointing device, the now ubiquitous mouse. The software programmers of Excel took advantage of these features unique to the Macintosh by incorporating into Excel design elements such as pulldown menus and point-and-click data entry. Microsoft released its first version of Windows in 1987 along with an early version of Excel for Windows. In 1989, Version 3.0 of Windows was released along with a much-improved version of Excel, which was the only Windows spreadsheet until 1992. Lotus 1-2-3 and other spreadsheet applications ran on the MS-DOS operating system.

Over the course of the mid- to late-1990s, there were three major spreadsheets available for the Windows operating system: Microsoft Excel, Lotus 1-2-3, and Quattro Pro. Each of these spreadsheets was incorporated into an office suite of software by its respective publisher. Software office suites usually contain a word processor, spreadsheet, presentation software, and sometimes a database. Lotus 1-2-3 was purchased in 1995 (Power, 2000) by IBM and included in its Lotus SmartSuite software package. Quattro Pro became part of Corel's WordPerfect Office suite. Excel, which has long been a part of the Microsoft Office suite, became the dominant spreadsheet on the market, accounting for about 90 percent of all spreadsheet sales (Krazit, 2002; Walkenbach, 2003).

More recently, low-cost and open source office suites have become available over the web that include spreadsheets. Open source software is developed collectively by volunteers in the worldwide programming community and made freely available to the public. Sun Microsystems' StarOffice 7.0 is available at no cost to academic and research institutions and is very competitively priced for businesses (www.sun.com/software/star/staroffice/index.xml). OpenOffice is an open source, feature-rich office suite that contains Calc, a fully functional spreadsheet. It is available at www.openoffice.org/product. The spreadsheets in these two office suites contain the full array of common spreadsheet tools; can read, write, and export Microsoft Excel files; and represent viable alternatives for individuals and agencies seeking lower-cost options in spreadsheet software.

The brief history of personal computer spreadsheets evidences three noteworthy trends. First, since the development of VisiCalc, spreadsheets have continued to expand the array of tools and features available to users. Spreadsheet users are no longer limited to simply calculating formulas for columns and rows of numbers. They can import data from multiple sources including databases and other types of spreadsheets. They can analyze complex data sets with descriptive and inferential statistics and then graphically represent the results with the charting tools of spreadsheets. Moreover, the tables, charts, and diagrams that users can now create with modern-day spreadsheets can be readily exported to other tools of electronic communication, including word processing, presentation, and web-authoring software (Patterson, 2000).

This integration of spreadsheets with other types of software is the second noteworthy trend in the history of spreadsheets. Early spreadsheets could generally produce printed output, but had no capacity to interact with other forms of software in sharing information. This electronic isolation was partially a function of the limitations of the MS-DOS and early Windows operating systems. The barrier between word processing and spreadsheets was broken by the release of Microsoft Word and Excel for the Apple Macintosh, which allowed for the cutting and pasting of charts and tables from Excel to Word. Since then, spreadsheet publishers have continued to expand the range of data and graphics exchange capabilities of their respective products. Most spreadsheets can now import data in a wide range of spreadsheet, database, and text file formats and export or share data, summary tables,

charts, and graphics to word processors, databases, and web pages, and many other applications. These capabilities are described in greater depth in subsequent chapters.

The third trend evident over the history of spreadsheets is their increasingly widespread utilization. Though it is difficult to measure directly the change over time in the use of spreadsheets by either professionals or in the general population, several facts are suggestive of their increased utilization. A search of www.amazon.com database using "spreadsheet" as the search term found listings for 3,058 available books. An identical search of www.barnesandnoble.com resulted in a list of 3,373 books related to spreadsheets. A search of Business Source Premier, an electronic database of business publications, found 4,253 articles that included the terms "spreadsheet" or "spreadsheets." A search of PsychInfo, an electronic database of journal articles, dissertations, and books related to psychology, found 177 listings of sources using the terms "spreadsheet" or "spreadsheets." An identical search of Social Sciences Abstracts, a multidisciplinary index of journals in the social sciences, found 64 articles. Rather remarkably, a search of Social Work Abstracts found only four articles using either of the two search terms. A search of www.google.com using the term "spreadsheet" found 1,970,000 web pages in which the word appeared. When "social work" was added to the Google "spreadsheet" search, 8,600 web pages were found. A mere 0.004 percent of all references to spreadsheets on the World Wide Web also include content somehow related to social work. In contrast, if the term "psychology" is added to "spreadsheet" on a Google search, 47,000 web pages are located. A search of World Wide Web Resources for Social Workers (www.nyu.edu/socialwork/wwwrsw) found only three articles using the search term "spreadsheet," none of which was directly related to social work.

It seems safe to conclude this brief history of spreadsheets with two observations. First, the sheer volume of spreadsheet-related publications, whether electronic or hard copy, is indicative of the widespread interest in their use and applicability to a broad spectrum of data analysis and information dissemination tasks. Second, if the dearth of academic literature and low volume of web pages are valid indicators, then it appears that social workers have yet to fully appreciate the utility of spreadsheets and take them up as tools of social work practice. It is this unfortunate neglect of a highly versatile tool that we seek to redress with this text.

Why Use Spreadsheets for Data Analysis in the Social Services?

In order to address the robust potential of spreadsheets as data analysis tools in the social services, it is perhaps necessary to first address the question of why social service professionals should concern themselves with "data analysis" in the first place. In the opening paragraph of this chapter we describe how information management, manifest in the collection, analysis, and subsequent communication of findings or information, is becoming a required skill set for social service professionals.

The necessity of social service information management skills is driven by three key factors: professional and fiscal accountability, ethics, and decision making (Montcalm & Royse, 2002). Social service agencies and professionals are increasingly being called on by funding sources, whether local, state, federal, or private insurers, to demonstrate both the actual delivery of services and efficacy of the services provided. In other words, it has become increasingly important to measure the outcomes of services provided. The accountability expectations of funding sources necessitates the collection of service delivery information, its subsequent analysis and evaluation, and the lucid communication of results.

For social workers, the confluence of ethics and social services information management skills arises from their Code of Ethics. The 1999 revision of the National Association of Social Workers (NASW) Code of Ethics states in section 5.02, Evaluation and Research (2003):

> (a) Social workers should monitor and evaluate policies, the implementation of programs, and practice interventions.
>
> (b) Social workers should promote and facilitate evaluation and research to contribute to the development of knowledge.

The ethical imperative conveyed in this section of the Code of Ethics is unequivocal. Social workers are expected to "monitor and evaluate" and "promote and facilitate evaluation and research." The intersection of this ethical obligation and information management is borne of pragmatism. The most efficient way to monitor and evaluate practice and facilitate research is with the employment of information technology.

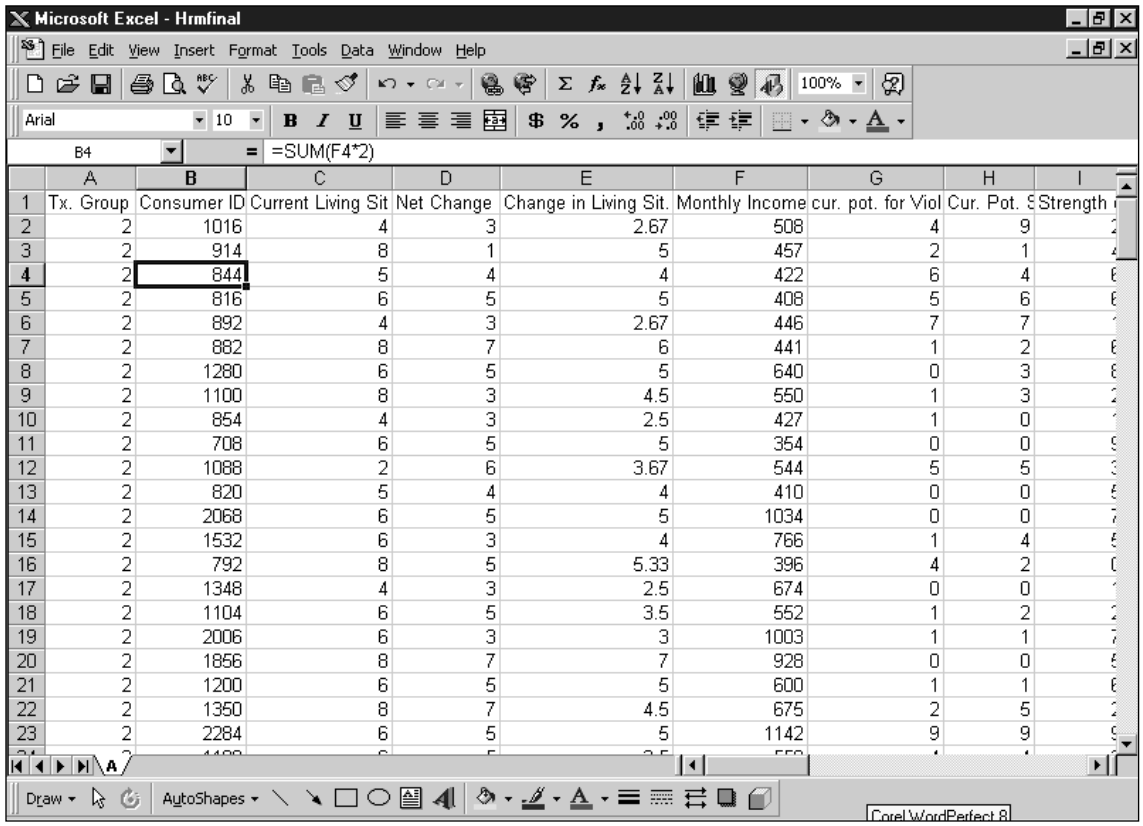
The third factor driving the need for information management skills in the social services is the ever-present need to make decisions (Montcalm & Royse, 2002). Social workers and social services professionals are called on daily to make critical decisions on matters including child protection, staff allocation for service delivery, agency fiscal resources utilization, and governmental policy recommendations, just to name a few. There are of course many ways to make a decision. Caprice, gut hunches, and intuition are means to arrive at a decision, albeit perhaps an ill-informed and intellectually unsupportable one. Proctor (2002) states, "Decision making in social work is high-stakes work" (p. 3). She asserts as well that decision quality can be impaired by the lack of information, as well as other factors. All too often, social service agencies have a wealth of information that is collected and stored on paper, an unfortunate decision that amplifies the complexity of information retrieval and synthesis. The point here is that decision making in the social services can and should be improved by information management skills that include electronic data collection, analysis/synthesis, and reporting. As we will see, social service practitioners can employ spreadsheets in each of these three domains of information management.

One of the great advantages offered to social service practitioners by spreadsheets is that they are commonly available on most personal computers. Today they are almost always part of an office suite of software applications. Sometimes the office suite is included on the computer as part of the original software and in other instances the office suite is purchased separately and loaded onto the computer. As described previously, there are now low-cost and open source office suites that are downloadable from the World Wide Web. It is now increasingly difficult to find a personal computer without some form of spreadsheet software on it. Consequently, social service practitioners and agencies have the readily available potential to employ spreadsheets for a spectrum of information management tasks.

Spreadsheet Basics

The overarching purpose of this book is the demonstration of spreadsheets' flexibility and data analytic power as tools of practice. In preparation to do that, it may be helpful to review the basic elements of spreadsheets. A spreadsheet is essentially a table composed of rows, columns, and cells. A column is a vertical line of boxes with a letter identifying each column (maranGraphics, 1996). A row is a horizontal line of boxes with a number identifying each row. A cell is a single box in the spreadsheet, which is the intersection of a row and a column. The cell reference is the address of the cell, which is composed of the column letter and the row number. For instance, cell B4 is located in column B on row 4. Figure 1.1 shows a spreadsheet in which cell B4 is highlighted (see the accompanying CD-ROM video animation: "Spreadsheet Tour").

Three types of information may be entered into a cell: labels, values, and formulas. Labels are explanatory text such as the name of a variable that appears at the top of a column and identifies the information contained in the column. Values are the data, both numerical and text, that are collected in the spreadsheet. Formulas perform calculations,



	A	B	C	D	E	F	G	H	I
1	Tx. Group	Consumer ID	Current Living Sit	Net Change	Change in Living Sit.	Monthly Income	cur. pot. for Viol	Cur. Pot. S	Strength
2	2	1016	4	3	2.67	508	4	9	
3	2	914	8	1	5	457	2	1	
4	2	844	5	4	4	422	6	4	
5	2	816	6	5	5	408	5	6	
6	2	892	4	3	2.67	446	7	7	
7	2	882	8	7	6	441	1	2	
8	2	1280	6	5	5	640	0	3	
9	2	1100	8	3	4.5	550	1	3	
10	2	854	4	3	2.5	427	1	0	
11	2	708	6	5	5	354	0	0	
12	2	1088	2	6	3.67	544	5	5	
13	2	820	5	4	4	410	0	0	
14	2	2068	6	5	5	1034	0	0	
15	2	1532	6	3	4	766	1	4	
16	2	792	8	5	5.33	396	4	2	
17	2	1348	4	3	2.5	674	0	0	
18	2	1104	6	5	3.5	552	1	2	
19	2	2006	6	3	3	1003	1	1	
20	2	1856	8	7	7	928	0	0	
21	2	1200	6	5	5	600	1	1	
22	2	1350	8	7	4.5	675	2	5	
23	2	2284	6	5	5	1142	9	9	

FIGURE 1.1 Spreadsheet of Client Data

such as the average or sum of a column or row of values. Formulas are one of the means through which data in a spreadsheet is analyzed. Formulas make it possible to test a range of "what if" scenarios. For instance, in planning a budget, a manager might examine the impact of a 2 percent, 4 percent, or 5 percent salary increase for staff from her available annual resources. The formula would automatically recalculate her personnel budget as she multiplied current salary cost by the possible 2 percent, 4 percent, or 5 percent salary increase. See Table 1.1 for an example. A caseworker also might help a client plan a household budget through exploring a range of "what if" scenarios around spending for the month. This capacity to test scenarios makes the spreadsheet a valuable tool in testing financial options in both administration and clinical practice settings.

TABLE 1.1 Spreadsheet for Salary Increases

Employee	Present Salary	2% Increase	4% Increase	5% Increase
J. P. Dole	32,000	32,640	33,280	33,600
Q. A. Smith	23,500	23,970	24,440	24,675
U. R. Wright	19,800	20,196	20,592	20,790
D. A. Day	18,000	18,360	18,720	18,900
Total	93,300	95,166	97,032	97,965
Net Increase in Personnel Cost		1,866	3,732	4,665

The following features are commonly found in spreadsheets.

1. Data Import/Export—The importation/exportation of data to and from a variety of sources, including database tables.
2. Functions—Built-in, predefined formulas for deriving specific mathematical, logical, informational, or statistical results.
3. Graphing—Data in the spreadsheet visually represented in a range of graph and chart formats, such as bar charts and area graphs.
4. Formatting—Options in formatting of cell contents including bold, italics, underline, justification (right, left, center), number representation (percentage, scientific, decimal places, etc.), row and column height and width, and specification of cell border styles.
5. Data Management—Information in spreadsheets sorted in a variety of ways, analyzed and summarized in pivot (cross-tabulation) tables, selected based on specified filters (e.g., all men over 50), and grouped in a variety of ways.
6. Statistical Analysis Tools—A collection of statistical procedures ranging from simple descriptive statistics, frequency counts, and histograms to inferential statistics.
7. Drawing Tools—A range of graphics tools available to enhance visually spreadsheet documents including lines, arrows, shapes, colors, and graphically augmented text tools, which add shape, color, and dimension to text.
8. Multidimensionality—Most spreadsheet programs create workbooks of spreadsheets (a.k.a. worksheets) that allow the linkage of a series of spreadsheets by formulas. The result is that a change in one spreadsheet will affect all the other spreadsheets to which it is connected by a formula. For instance, an administrator could create a spreadsheet with the agency's budget on it. This spreadsheet would be linked to three other spreadsheets in the same workbook, each spreadsheet representing a department's budget. The four spreadsheets might be linked by a formula that allocates a percentage of the agency's personnel budget to each of the three departments. Any change in the spreadsheet cell that contains the agency's budget would alter, in turn, the budget allocation for each department.

Data In

The aggregation of data into a spreadsheet can occur through many routes. Data may be directly input by typing it into the cells of the sheet. Many spreadsheets have the option of creating a data entry form, making it possible to enter data in much the same manner as one would enter data into a database.

Most spreadsheets allow for the importing of data files from databases. This feature makes it possible to pull data drawn from a database into a spreadsheet in order to analyze and summarize it in ways that are not possible with most databases. This data importation from databases to spreadsheets is an extremely useful option. Many progressive or better-funded social service agencies now have databases to track client demographics, services delivered, and services outcomes data. These databases can be queried for information of interest. A query is the extraction of data from database tables based on specifically defined criteria. For instance, if one wished to understand what services are delivered to a subset of the population served by an agency (e.g., all clients living in a particular geographic region), then one might request from the agency's database administrator a file containing a list of all clients from that geographic region along with demographic information and a listing of the services they received. Once the database query is run, the resulting file, often called a "flat file," could be imported into a spreadsheet for an analysis of services delivered by a demographic group. The specific steps of such an analysis are presented in Chapter 9.

Using Formulas and Functions

Regardless of the means by which data arrives in a spreadsheet, whether by direct data entry, a data entry form, or importation from a database, formulas and functions are commonly the

TABLE 1.2 Simple Formula Example

	A	B	C
1		45	53
2		56	76
3		75	15
	Sum	176	=sum(C1:C3)

first tools used to understand the data. Formulas and functions enable spreadsheet users to conduct data analysis procedures ranging from the simple to the dreadfully complex. For now, we will limit our introduction of formulas and functions to the simple end of the continuum.

Simple Formulas. Formulas define the calculations the spreadsheet is to perform (Parsons, Oja, & Auer, 1995). A formula is initiated by clicking on the cell in which the results of the calculation is to appear. Every formula has an operation symbol indicating to the spreadsheet the presence of a formula. In Microsoft Excel, the operation symbol that appears at the beginning of any formula is the equal sign (=), whereas Lotus 1-2-3 and Corel Quattro Pro use the "at" symbol (@). For instance, this formula in an Excel spreadsheet tells the application to add the numbers contained in the formula =45+56+75. In Lotus 1-2-3 and Quattro Pro the formula would be @(45+56+75). Once the values are entered into the formula, pressing the Enter key returns the numerical result of the formula's calculations. Likewise, formulas can use values in other cells by entering their cell references into the formula. For example, in Table 1.2 the formula summing the three numbers could be expressed as =(B1+B2+B3) or as =(B1:B3). Column C contains a set of values that will be added with the formula in cell C4. See the accompanying CD-ROM video animation: *Formulas and Functions.*

Operators in Formulas. The formulas in the preceding paragraph contain several symbols, including =, +, (, and). These symbols are referred to as operators. Operators are used in formulas for arithmetic calculations, to reference cells, to link portions of equations, and to segment portions of equations. Table 1.3 contains a list of commonly used operators. Technically speaking, parentheses are not operators but instead indicate in the formula the order of operation. When calculating a formula, the spreadsheet will first calculate the portion of the formula contained within parentheses and then use the results to calculate the remainder of the formula. For instance, in the equation =(5+7+9*2), the formula adds 5+7 and then adds 18, the results of 9*2, to produce a product of 30. Whereas, if the equation is written as =(5+7+9)*2, the sum of the numbers within the parentheses is 21, which is then multiplied by 2 for a product of 42 (Dretzke & Heilman, 1998). Basic algebraic order of operations procedures are applied in spreadsheet formulas. This means that operations in

TABLE 1.3 Operators for Formulas

Symbol	Meaning	Formula Example
=	Equal—Used at the start of every equation	=(53+76+15)
+	Plus sign—Used for addition	=(C1+C2+C3)
–	Subtraction or a negative number	=(76–15)
*	Asterisk—Used for multiplication	=(C1*C2) or =(53*76)
/	Forward slash—Used for division	=(76/15) or =(C2/C3)
%	Percent	=(76*10%)
^	Caret—Used for exponentiation, e.g., to square or cube a number	=(76^2) or =(76*76) =(76^3) or =(76*76*76)
:	Colon—Used to indicate a range of cells	=sum(C1:C3)
,	Comma—Used to combine two or more references	=sum(B1:B3,C1:C3)
( )	Parentheses—Used to segment portions of an equation and to specify the order of operation	=sum(B1:B3)*(C1^2)

parentheses are calculated first, then exponential expressions (10^2), next division and multiplication operations, and finally addition and subtraction.

Functions. Some spreadsheet calculations are most readily accomplished with functions, instead of writing a formula or equation. "A 'function' is a special prewritten formula that provides a shortcut for commonly used calculations" (Parsons, Oja, & Auer, 1995, p. 32). Essentially, functions are mathematical or logical procedures that are inserted in one or more cells and perform a calculation or logical operation. Examples of functions include statistical functions such as average, standard deviation, random, median, and logical functions such as "if," "and," and "true." In Excel, functions are found in the menu bar under Insert, Function. When Function is selected from the menu bar, a pop-up window appears with categories of functions listed on the left and function names listed on the right. Categories of functions include Math & Trig, Logical, Financial, and Statistical. For our purposes, Statistical is the category of function we will most commonly employ. It offers a wide range of statistical formulas from which we will use a limited number. The great utility of functions is that they allow one to quickly and accurately perform simple to complex calculations and analyses on data. Functions also reduce formula typing errors. Further, each function provides step-by-step information and instructions necessary to employ the selected function for its intended purpose (Black, 1999). We will use functions for a number of data analysis purposes in this book.

Advanced Tools for Understanding Data

In addition to formulas and functions, spreadsheets offer a number of other data analysis tools. Each of these data analysis tools is demonstrated in subsequent chapters, but in the service of this overview of the versatility of spreadsheets, some of them are described in Table 1.4. Taken together, this collection of data analysis tools makes it possible to conduct a spectrum of data analysis tasks ranging from simple data sorting to complex inferential statistical analysis.

One of the tools in Table 1.4 is both (for our purposes) unfortunately named and particularly important to understanding the power and versatility of spreadsheets as tools of social service practice, particularly in data analysis. This collection of statistical procedures is rather indistinctly named Data Analysis Tools in Excel. Microsoft's use of this poorly differentiating label for Excel's statistical tools obscures both the data analytic power of the other tools listed in Table 1.4 and the statistical utility of the collection of procedures found in the Data Analysis Tools.

TABLE 1.4 Data Analysis Tools of MS Excel

Chart Wizard	A step-by-step guide for the creation of charts. Enables the graphical representation and exploration of data.
Data Analysis Tools	A collection of tools for the organization, statistical analysis, and interpretation of numeric data.
AutoFilter	A tool for filtering data sets for subsets of data based on selected criteria from the data set.
Functions	A predefined formula to perform a calculation on a specified range of values.
Pivot Table	A tool to collapse and summarize data in order to understand the relationships between variables. The means to create cross-tabulation tables.
Sort	A tool for sorting data sets or subsets of data in specified order (ascending or descending).

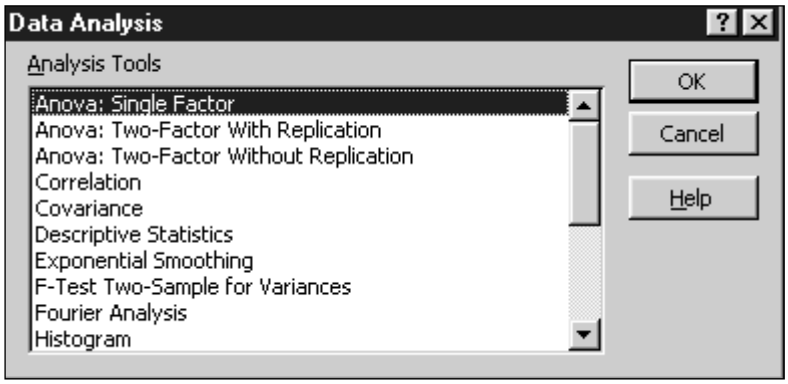


FIGURE 1.2 Data Analysis Tools

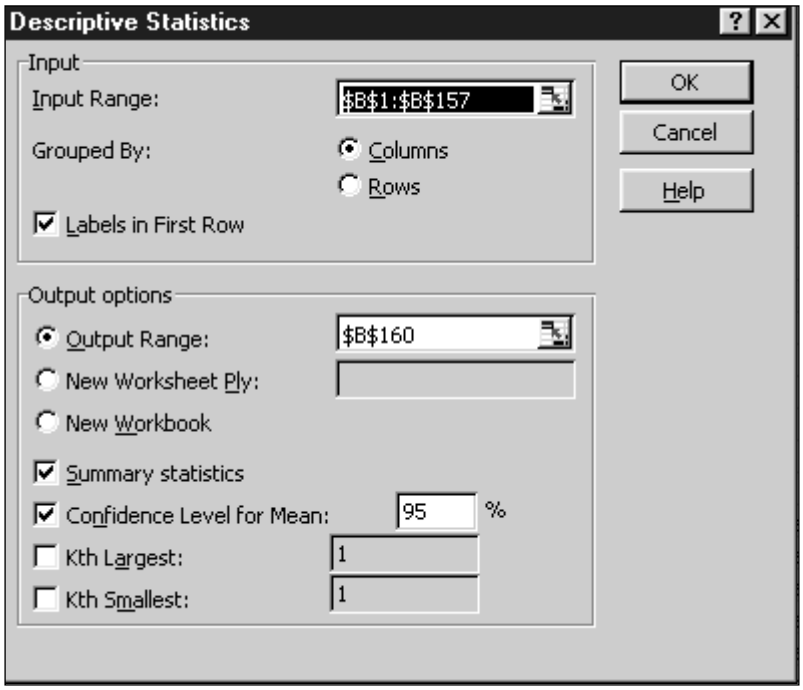


FIGURE 1.3 Data Analysis Tools—Descriptive Statistics

TABLE 1.5 Data Analysis Tools—Descriptive Statistics Output

AGE	
Mean	62.99
Standard Error	0.75
Median	64.00
Mode	70.00
Standard Deviation	9.40
Sample Variance	88.37
Kurtosis	1.68
Skewness	–0.44
Range	64.00
Minimum	23.00
Maximum	87.00
Sum	9827.00
Count	156.00
Confidence Level (95.0%)	1.49

The Data Analysis Tools are found in Excel's menu bar Tools, Data Analysis. If Data Analysis does not appear under the Tools menu, then select Add-Ins. The Add-Ins window will appear; in that window, select Analysis ToolPak, then click OK. This will add the Data Analysis Tools to the Tools menu. Figures 1.2 and 1.3 display the Data Analysis Tools window and the Descriptive Statistics dialog box. Table 1.5 contains the statistical results produced by the Descriptive Statistics tool.

Though Figure 1.2 does not show the entire selection of statistical procedures available in Data Analysis Tools, it in fact contains a robust collection of statistical procedures. These statistical procedures are widely used in business, scientific, research, and financial settings. As will be demonstrated in subsequent chapters, they are more than sufficient to address the data analytic requirements of most social service professionals and agencies.

A Comparison of Spreadsheets to Statistical Analysis Software

The alternative data analysis tool to the spreadsheet is statistical analysis software. There are many brands of statistical analysis software, including SPSS, STATISTICA, SYSTAT, all of which are commercially available products. There are also a number of statistical software packages available in the public domain at no or low cost. See http://members.aol.com/johnp71/javasta2.html#General for a list of no- or low-cost packages. There is, of course, a great deal of variability in the appearance of, functionality of, and statistical producers included in statistical software applications. Some of the commonalities among most statistical software packages include a graphical user interface, ability to create and edit data files, import/export data tools, charting tools, pivot tables (cross-tabulation), and a range of statistical procedures and tests.

There are many commonalities and some differences between spreadsheets and statistical analysis software. Spreadsheets and statistical analysis software are compared in Table 1.6.

TABLE 1.6 Spreadsheets Compared to Statistical Software

Assessment Category	Spreadsheets	Statistical Software
Ease of use	Ranges from easy to challenging depending on the task	Ranges from easy to challenging depending on the task
Graphical user interface	Yes	Yes for most packages
Versatility	Capable of multiple tasks	Statistical analysis/graphing
Direct data entry	Yes	Yes
Data entry form tool	Yes	May require add-on tool
Mathematical/statistical functions	Preformatted equations available for rapid calculations/data analysis tools	Simple to sophisticated set of statistical procedures selected from pulldown menus
Charting and graphing	Flexible graphing tools with easy customization	Graphing tools selected from pulldown menu. Formatting and customization more challenging
Data sorting/filtering	Yes	Yes
Pivot tables (cross-tabulations)	Yes	Yes
Diagramming and drawing tools	Yes	No
Importation of data from database	Yes	Yes
Web output (html file format)	Yes	Yes
Budget analysis	Yes	No
What-if scenarios[1]	Yes	No

[1]A data analysis procedure to test different outcomes based upon a range of input options. For instance, one might test how an agency's total budget would be impacted by 2%, 4%, 6%, or 8% raises for all employees. As each of the raise percentages are input into the spreadsheet, formulas recalculate the total budget.

Sources: Johnson, 2001; Nash & Quon, 1996.

The Paradox of Spreadsheets and Statistical Software in Social Service Practice

It is quite common in social work and other human and social services professional education programs for students to be taught to use statistical software such as SPSS. SPSS is a powerful statistical software package that is widely used in undergraduate and graduate education. Faculty members, especially those who teach research and statistical classes, often have some level of proficiency with SPSS because they likely had to use it in the data analysis required for their doctoral research. The cycle of learning pattern is that faculty learn SPSS or some other statistical software package in their doctoral education and then they teach the use of the same software in their undergraduate and graduate classes. The question then arises, Why should social workers and other social service professionals learn spreadsheets for data analysis when SPSS works so well and so many professors teach its use? This question is perhaps best answered by a simple bit of research. If you work in a social service agency or if you are a student doing a field placement or internship in a social service agency, examine the available programs on any computer to which you have approved access. See if you can find a spreadsheet or a statistical software package on the machines available to you. If the machines you are examining have the Windows operating system, then under the Start, select Programs. If you are using an Apple Macintosh, then click on the hard drive and look in the Applications folder. You may also want to ask a supervisor or information technology (IT) person in the agency whether computers in the agency have available spreadsheets or statistical software packages. Some supervisors may be better prepared or more technologically proficient and able to answer this question. Most agency IT staff should be able to offer a reasonably accurate answer.

The point of this small research project is that it is much more likely that you will find a spreadsheet on the social service agency computers you examine than a statistical software package. There are a number of good reasons for this likely fact. First, as described previously, spreadsheets are commonly installed on personal computers either by the vendor, the agency, or the personal owner, usually as part of an office suite. Second, the intended function of statistical software packages is too specific for them to be broadly installed on personal computers. The third factor is that widely taught statistical software packages such as SPSS are relatively expensive. There are less expensive, but perhaps less robust and powerful, statistical software packages available. However, a careful examination of social work research texts we conducted indicates that SPSS is the most commonly referenced statistical package (Patterson, Basham, & DeCoster, in preparation). It is our hypothesis that within social work, the use of statistical packages such as SPSS for data analysis is widely taught in undergraduate and graduate programs; paradoxically, students and professionals are far more likely to have access in social service practice settings to spreadsheets for data analysis than statistical packages. The unfortunate likely consequence of this mismatch between education and practice is that professionals are less likely to engage in practice and program evaluation and research. We believe practice and program evaluation are critical to the vitality of the profession. The purpose of this book is to prepare students and professionals to conduct the data analysis necessary to produce high-quality practice and program evaluation with the tools most readily available to them: spreadsheets.

Summary

This chapter provides an introduction to spreadsheets and their potential as a tool of practice in the social services. The chapter describes the history of the development of spreadsheets in their pre-electronic and electronic forms. Three trends in the development of spreadsheets are noted: the expansion of tools and features, improving ability to import and export information in multiple forms and formats, and increasingly widespread utilization. This chapter presents a rationale for the greater adoption of spreadsheets in social service

practice. An overview of the basic features, functions, and data analysis tools of spreadsheets is presented. Finally, this chapter provides a comparison between spreadsheets with statistical software followed by an examination of spreadsheets' paradoxical underutilization in the social services despite their widespread availability.

REVIEW QUESTIONS

1. Who created the first personal computer-based spreadsheet for what purpose, and what was it called?
2. What are the three noteworthy trends in the history of personal computer spreadsheets?
3. How can the ethical imperatives to "monitor and evaluate" and "promote and facilitate evaluation and research" be accomplished with spreadsheets?
4. What types of information can be entered into spreadsheet cells?
5. What does the term "multidimensionality" mean in reference to spreadsheets?
6. What do formulas and functions do in spreadsheets?
7. How are the Data Analysis Tools activated in Excel and, once activated, where are they found in the menu bar?
8. What has been the historical paradox of spreadsheets in social service education and practice?