

APPENDIX B: COMPUTER SOFTWARE TO COMPILE AND ANALYZE DATA

APPENDIX B: COMPUTER SOFTWARE TO COMPILE AND ANALYZE DATA

I. Spreadsheet Software

What is a Spreadsheet?

A spreadsheet consists of a **grid of rows and columns**. Each box in this grid is called a **cell**.

	"Column" A	B	C	D	E	F	G
"Row" 1	"Cell"						
2							
3							
4							

Spreadsheets can be created by hand, using paper and pencil, or created with a computer software program. In computer software programs, the columns are labeled with letters (A, B, C ...) and the rows are numbered (1, 2, 3 ...). Thus any box or cell in this grid can be referred to in terms of its column and row label, for example C3, D17, or AA45.

Each cell in this grid can contain numbers, percentages, text, or formulas. The size of each cell can be adjusted to fit its contents. This is no problem with numbers, percentages and formulas. However, the amount of text that can be typed into each cell of a spreadsheet depends on how big you want your spreadsheet to be once you're done. Since the benefit of a spreadsheet is to be able to "eyeball" your data, it makes sense for as much of it to fit onto as few pages as possible. You can limit the amount of text entered into each cell by summarizing text points or using short-hand terms that are explained elsewhere.

Best use of spreadsheet software

Spreadsheet software is a good way to enter and maintain lots of collected data. Without a spreadsheet, you have to sort through stacks of surveys or notes every time you want to see your data, and calculate your summary statistics by hand. Once you enter your data into a spreadsheet, it is easily accessible and can be analyzed in different ways.

The essential quality of using computer software instead of paper and pencil for your spreadsheet is that you can use formulas. These formulas can relate the numbers in

one cell to others in the spreadsheet. For example, if “5” is entered in cell B12, and “6” is entered in cell D15, and the formula B12+D15 is entered in cell F20, then cell F20 will display the value “11”. The types of formulas most computer software offer include most mathematical functions. These are often called **summary statistics**, and usually include:

- The addition of two or more cells
- The subtraction of two or more cells
- The multiplication of two or more cells
- The division of two or more cells
- The mean average of the data in a range of cells
- The count of the numbers of responses in a range of cells
- The sum total of the data in a range of cells

Spreadsheet software can also be used to create **tables and charts** for presenting your data. All spreadsheet software packages include formatting for creating printable tables—including gridlines, shading, font options, and colors. They also include graph wizards for selecting rows and columns of data in your spreadsheet to create different kinds of graphs, such as line, bar, and pie graphs, and include different formatting options regarding graph titles, x-and y-axis parameters and titles, fonts, colors and shading.

What are some disadvantages of spreadsheet software?

When using spreadsheet software to enter and analyze your data, remember that the software can't tell you what kind of summary statistics, charts or graphs would be best to use with your data. Nor will it tell you how to interpret them. You need to be aware of what kinds of formulas you are creating and whether or not the “answer” the software creates for you is correct and meaningful.

Spreadsheets are best for data that can be somewhat quantifiable. Responses to data collection methods that take the form of long anecdotes or discussions are difficult to fit into a spreadsheet. Your best bet in that situation is to use word processing software, and only use a spreadsheet software program to create charts that help you characterize qualitative themes or participant characteristics.

What spreadsheet software programs are available?

The most common spreadsheet software package is Microsoft Excel. It is available for both PCs and MACs. The benefit of Excel is that it is compatible with Microsoft Word and Microsoft Power Point, which allows you to: 1) copy and paste text between a document, spreadsheet and slide presentation; and 2) copy graphs and charts from a

spreadsheet into a report or slide presentation. It is also the most widely available of the spreadsheet software, and so will be available on most computers being used by your community partners, as well as computers accessible for free use in local libraries or computer resource centers.

Other spreadsheet software that is similar to Excel:

- IBM's Lotus 1-2-3
- Mariner Calc
- Zim FastFeedback

Where can I go for more information?

- Analyse It, a statistical software add-in for Microsoft Excel (download a free 30-day trial version: <http://www.analyse-it.com/default.asp>). This software complements Excel by adding on higher-level statistical analysis functionality to the existing Excel program.
- Florida Gulf Coast University, Excel 2000 Tutorial: <http://www.fgcu.edu/support/office2000/excel/>
- The Land Grant, "Learning Lotus 1-2-3": <http://www.lgta.org/123/123.contents.htm>
- Mathsnet, "The A to Z of Spreadsheets": <http://www.mathsnet.net/a2zofssheets.html>
- Microsoft Excel webpage: <http://office.microsoft.com/home/office.aspx?assetid=FX01085800>
- University of South Dakota, "Excel Tutorial": <http://www.usd.edu/trio/tut/excel/index.html>
- Computing and Communications Services Office, "Lotus 1-2-3 the Basics": http://w3.aces.uiuc.edu/AIM/CCSOcourses/new-lotusbasics95_final.html
- ZDNet, Reviews and Prices of Spreadsheet Software: http://reviews-zdnet.com.com/Business_software/4502-3527_16-0.html?tag=dir-ss
- Microsoft Access or Filemaker Pro searchable Help Contents
- Do your own search by typing "Microsoft Excel" or "spreadsheet software" into the Google search engine: <http://www.google.com>

2. Database Software

What is a database?

A database is a collection of data or information that is related—such as name, address, phone number of a respondent. Databases are often called "relational databases" or "relational database management systems". This relational functionality is what makes

databases different from spreadsheets—although data is entered on a screen that looks just like a spreadsheet, the data is structured and stored in a way that can be easily sorted and retrieved in many different forms. You may already encounter database software every day in your work, as databases are more often used for management purposes (such as billing, medical records, and client tracking) than for research or evaluation.

What is the best use of database software?

Database software can be used to enter and store data from your community assessment data collection methods, which can include quantitative as well as qualitative data. Database software can be customized, so that you enter only the types of data (“fields”) you need to maintain the database.

Once data is entered or imported into a database you can run “queries” or create “reports”. Queries are specific data retrievals that give you only the data you wish to look at. For example, you could ask to see the data for all community members who reported smoking more than a pack of cigarettes a day. Or, you might want to see the insurance status of all African-American teenage male respondents. Since queries are customizable, you can retrieve any data in any way you wish.

Reports show data from multiple records at once. These can be generated from your data tables, from queries, or in any way you wish to look at the data. Since reports are also customizable, you can select to see any data you want in any format you want. Looking at your data in different ways and in different report formats helps you to see relationships and analyze your data. Printing out reports on a regular basis also allows you to make sure that you are collecting the type of information through your community assessment that answers your primary questions. It also helps you create and submit reports to community partners, the community, and target audiences—like policymakers and funders.

What are some disadvantages of database software?

As with spreadsheets, it is important to understand that database software does not organize and analyze data for you. It merely is a data compilation tool that can be customized and changed throughout your data collection, data analysis and data report-writing phases without losing the basic relationships between pieces of data.

Learning to use a database software program can be more time-consuming than a spreadsheet software program. The extra features and customizability make the software more complicated, requiring a longer learning process. This must be understood up front if you decide to use a database software program to manage your community assessment data, and have never used one before.

What database software programs are available?

Again, the most commonly found database software is Microsoft Access, as it can be found - along with Microsoft Word, Excel and Power Point - in any Microsoft Office Professional suite of software. The benefit of using Access is that it makes for easier importing of data and text into it from Microsoft Word and Excel, exporting of tables into Word, Excel and Power Point, and merging with Word for mailings or the like.

However, if you don't currently have Microsoft Office, then there are other database software programs that are just as good and may be significantly cheaper to buy. These include Claris Filemaker Pro and Microsoft FoxPro.

In addition, many clinics and hospitals have their own database systems that have been customized for their unique patient data collection and billing needs. If you are requesting anonymous and aggregated patient data from these sources, you may be able to receive data reports from these databases that help fulfill some or all of your data needs without having to collect data and create your own database.

Where can I go for more information?

Here are some helpful on-line resources:

- CompuMentor, Database consulting for nonprofits:
<http://www.compumentor.org/consult/db/default.html>
- Coyote Communications Technology Tip Sheet, Database, software and technology use: <http://www.coyotecom.com/database/index.html>
- Filemaker Pro website: <http://www.filemaker.com/>
- *Inside Microsoft Access*, the book: <http://www.elementkjournals.com/ima/>
- Microsoft Access website:
<http://office.microsoft.com/home/office.aspx?assetid=FX01085791&CTT=6&Origin=ES790020011033>
- Microsoft Access Tutorials:
<http://mis.bus.sfu.ca/tutorials/MSAccess/tutorials.html>
- The Access Web: <http://www.mvps.org/access/>
- Microsoft Access searchable Help Contents
- Do your own search by typing "Microsoft Access" or "database software" into the Google search engine: <http://www.google.com>

3. Statistical analysis software

What is statistical analysis software?

Statistical analysis software is typically used for the statistical analyses that spreadsheet and database software on their own cannot perform. As a result, most statistical analysis software programs are compatible with standard spreadsheet and database programs. After importing data from these programs into a statistical analysis program, they are able to calculate higher-level statistical analyses than either spreadsheets or databases.

Best use of statistical analysis software

Statistical analysis software can perform data analyses that range from basic descriptive tables (such as the number of clinic users by gender) to advanced statistics (such as regression analysis). In addition, these software programs also make it easier to calculate means, percentages, frequencies, and sums without having to create formulas or tables by hand. Some, like EpiInfo, also double as a database with data entry features, so that it can be the only software you need.

These programs also create data analysis output reports, which allow for easier transfer into report tables and charts for representing statistical analysis results. Some programs also create output and reports for direct import into word processing and web-based programs allowing for easier distribution.

What are some disadvantages of statistical analysis software?

Statistical analysis software can be extremely expensive, so be sure that you *absolutely* have to use it and that it meets your needs *well*, before investing in it.

Statistical analysis software is also difficult to use. Just as with spreadsheet or database software, you must keep in mind that these software programs are not able to determine which statistical analyses should be performed for your data. Nor are they able to interpret the analyses for you. Therefore, before using this software you should have a solid knowledge of statistics, how to perform the most appropriate analyses with your data, and how to interpret analytical results. Otherwise, the software will be unable to help you.

What statistical analysis software programs are available?

EpiInfo is a public domain software package designed by the U.S. Centers for Disease Control and Prevention (CDC) for the global community of public health practitioners and researchers, and is available for free download on their website. It provides for

easy database construction, data entry, and analysis with epidemiological statistics, maps, and graphs—so that all of the data collection, data entry, data management, data analysis and data reporting functions are all contained within one software package. The primary applications within EpiInfo are:

- **MakeView**—a program for creating forms and questionnaires which automatically creates a database.
- **Enter**—a program for using the forms and questionnaires created in MakeView to enter data into the database.
- **Analysis**—a program for producing statistical analyses of data, report output and graphs.
- **EpiMap**—a program for creating GIS maps and overlaying survey data on to them.
- **EpiReport**—a tool that allows the user to combine Analysis output, enter data—and any data contained in Microsoft Access or SQL Server—and present it in a professional format. The generated reports can be saved as HTML files for easy distribution or web publishing.

Although “EpiInfo” is a CDC trademark, the programs, documentation, and teaching materials are in the public domain and may be freely copied, distributed, and translated. For a tutorial on how to use this software, see <http://www.nhip.org/Learning%20Module/Epi2000/epiinfo2000.html>

There are several other statistical software packages that have a fairly user-friendly, Windows-driven design, and are well regarded among researchers. However, they usually require a high level of knowledge regarding statistics, data manipulation, and computer technologies. If your partnership or organization has the capacity and desire to use these more complicated software programs, you may want to check some of the websites and publications listed below. These programs include:

- SAS
- SPSS
- Stata
- XLStat

There are also several statistical analysis software programs especially designed for analyzing qualitative data. These include:

- ATLASTi
- Ethnograph
- HyperRESEARCH
- QSR NUDIST

- U.S. Centers for Disease Control and Prevention's free downloadable EZ-Text: <http://www.cdc.gov/hiv/software/ez-text.htm>

If you are considering using a software package for this purpose, please consult a research consultant or faculty member at a university who has conducted qualitative data analysis with such programs. These software packages can be very costly, and may end up being less useful than a simpler method, unless you are extremely experienced in both qualitative data analysis and computer software. For more information on this, please look on the Computer Assisted Qualitative Data Analysis Software Networking Project site: <http://caqdas.soc.surrey.ac.uk/index.htm>.

Where can I go for more information?

- Centers for Disease Control and Prevention, EpiInfo: <http://www.cdc.gov/epiinfo/> Contains free downloadable software, tutorials, and links to other web resources.
- B. Burt Gerstman, Data Analysis with EpiInfo: <http://www2.sjsu.edu/faculty/gerstman/EpiInfo/>
- Princeton University, Stata Tutorial: <http://www.princeton.edu/~erp/stata/main.html>
- SAS Software Customer Support Center: <http://support.sas.com/>
- SPSS Website: <http://www.spss.com/>
- Stata Statistical Software website: <http://www.stata.com/>
- Statpages.net, "Web pages that perform statistical calculations": <http://members.aol.com/johnp71/javastat.html>
- Statpages.net, "Free statistical software": <http://members.aol.com/johnp71/javasta2.html>
- UCLA Academic Technology Services, Resources to Help You Learn and Use Stata: <http://www.ats.ucla.edu/stat/stata/>
- XLState website: <http://www.kovcomp.co.uk/xlstat/index.html>
- Do your own search by typing the name of the statistical software program into the Google search engine: <http://www.google.com>

4. Computer Resources

If the organizations in your community partnership don't own or have access to computers, or computers with the software, Internet access, or other technological capacity you need, you may want to look for free computer resources in your area.

Here are some online resources you may want to check out:

- Alliance for Technology Access: Connecting children & adults with disabilities to technology tools: <http://www.ataccess.org/resources/>
- Asian Pacific American Network, Providing computer technical assistance for Asian/Pacific Islander community-based organizations in the Los Angeles area: <http://www.apanet.org/index.html>
- Center for Nonprofit Management: <http://www.cnmsocal.org/>
- Community Technology Foundation of California: Social justice, access & equity through computer technology: <http://zerodivide.org/>
- Computers in Our Future: <http://www.ciof.org/index.html>
- Techsoup: The Technology Place for Nonprofits: <http://techsoup.org/>