



Software Packages for Quantitative and Qualitative Research

Prof Oluwadiya KS

Provost, College of Medicine

Ekiti State University, Ado-Ekiti

www.oluwadiya.com

Definitions

Quantitative Research

- Quantitative methods emphasize objective measurements and the statistical, mathematical, or numerical analysis of data collected through polls, questionnaires, and surveys, or by manipulating pre-existing statistical data using computational techniques. Quantitative research focuses on gathering numerical data and generalizing it across groups of people or to explain a particular phenomenon

Qualitative Research

- Qualitative research is a type of research that collects and works with non-numerical data and that seeks to interpret meaning from these data that help us understand social life through the study of targeted populations or places.



Quantitative versus Qualitative Research

Quantitative	Qualitative
AIM: count & classify features, then construct statistical models to explain the observation.	AIM: complete, detailed description.
Recommended during latter phases of research projects.	Recommended during earlier phases of research projects.
All aspects of the study are carefully designed before data is collected.	The design emerges as the study unfolds.
Researcher uses tools, such as questionnaires to collect numerical data.	Researcher is the data gathering instrument.
Data is in the form of numbers and statistics.	Data is in the form of words, pictures or objects.
Objective - uses precise measurement & analysis of target concepts, e.g., questionnaires	Subjective – individuals' interpretation of events is important ,e.g., in-depth interviews etc.
Quantitative data is more efficient, able to test hypotheses, but may miss contextual detail.	Qualitative data is 'richer', time consuming, and less generalizable.
Researcher tends to remain objectively separated from the subject matter.	Researcher tends to become subjectively immersed in the subject matter.

Quantitative Versus Qualitative Software

Quantitative Software	Qualitative (CAQDAS)
Older History: dated back to the early days of computers	First programs started emerging in the early 80s
Widely employed by researchers in the field	More limited usage, many researchers are still employing traditional methods
Most packages do the same thing, only their approaches are slightly different, and their coverage are also similar	Vastly more different, no two qualitative programs take quite the same approach
Actually used to analyze the data	Merely serves as an aid to the analysis process. The researcher does the analysis

CAQDAS: Computer Assisted Qualitative Analysis Software

Research Software: A general look

- **Tools for Literature searches** ...Google Scholar, HINARI, PUBMED
- **Word Processors:** MS Word, Google Docs etc
- **Data Collection Software:** e.g. Open data Kit (ODK), and online survey instruments (e.g.Survey Monkey)
- **Reference Managers** e.g. Zotero, Endnote, Mendeley etc.
- **Archiving and Note Taking Software** e.g. MS OneNote,
- **Statistic software: Quantitative and Qualitative**
- **Cloud Storage:** Dropbox, MS One Drive, Google Drive, etc. Never lose data again

Statistics software can be divided into:

- On the basis of payment method
 - i. Free
 - ii. Shareware
 - iii. Paid
- On the basis of user interphase
 - i. Point and Click (Graphic User Interphase)
 - ii. Programming
 - iii. Combined
- On the basis of repository
 - i. On the Computer
 - ii. Online

1. Quantitative Analysis Software

Quantitative Statistics software: How many?

- There are many statistics software currently available
- AS of 09/09/2017, Wikipedia 58 software listed (https://en.wikipedia.org/wiki/Comparison_of_statistical_packages)
- These range from small niche packages to robust software that performs many analysis.
- **Payment Options**
 - Open source (Free): 19
 - Proprietary: 39
- **Operating system:**
 - Windows: 58
 - Mac OS: 37
 - Linus: 34



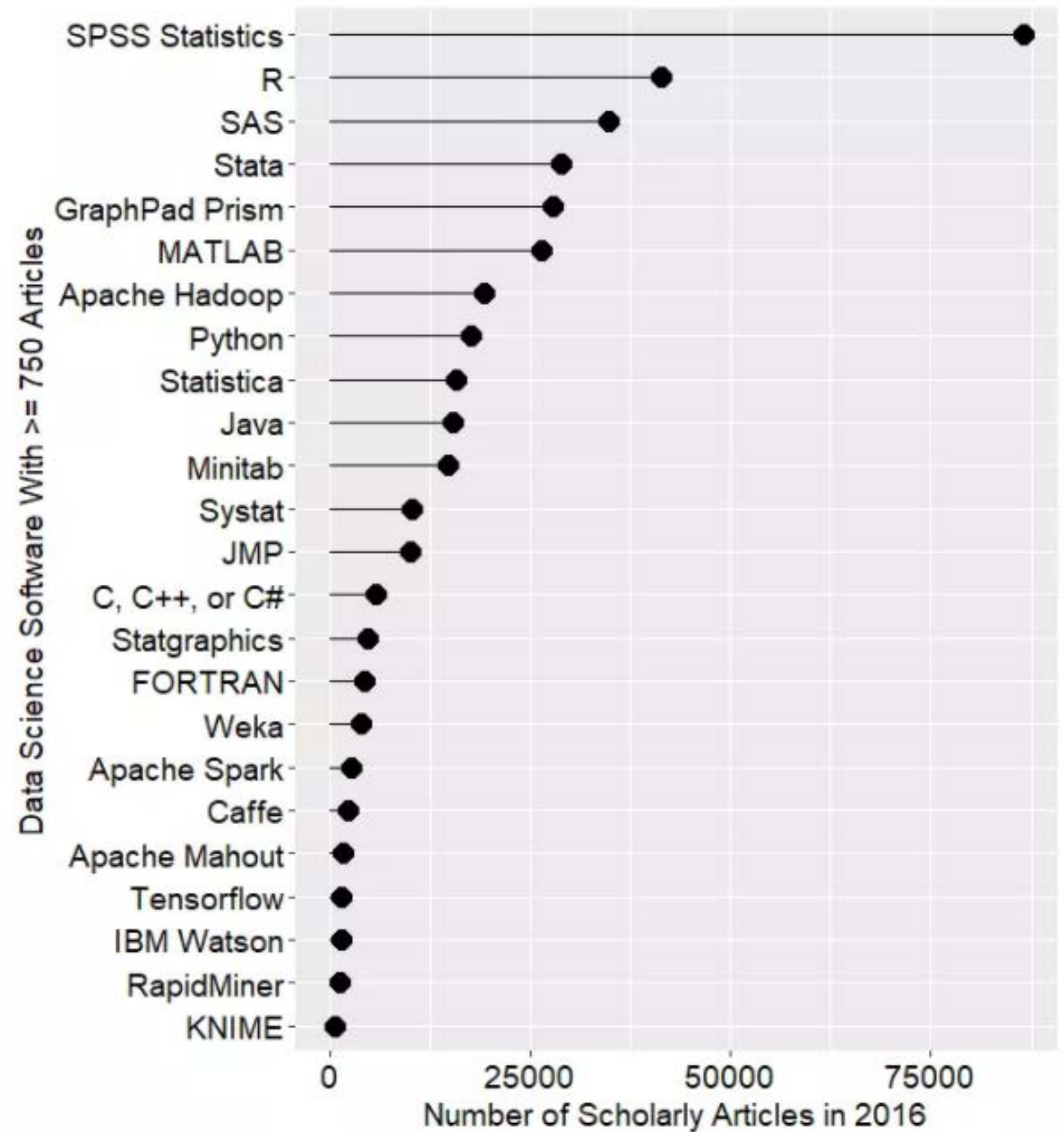
Key considerations for choosing a statistical package

- i. **Ease of use.....** SPSS and Minitab emerge as the best.
- ii. **Learning curve.....**SPSS the easiest to learn
- iii. **Depth of point and click procedures.....** SPSS is the clear leader on this criterion
- iv. **Range, quality, and ease of use of statistical procedures offered.....** R, SAS, and Stata, in that order, offer the widest ranges of statistical procedures followed by SPSS
- v. **Modifiability of analytical output specifications....**SPSS seems to be the best
- vi. **Range of graphical output offered....**R is the best, but learning curve is steep.

Key considerations for choosing a statistical package

- vii. Speed of handling large data sets.....SAS is the best.**
- viii. Ease and flexibility of data importation R has best data importation ability, followed by SAS, SPSS, and Stata**
- ix. Ease of results exportation....SPSS is by far the best**
- x. Thoroughness and interpretability of results outputNo clear leader**
- xi. Ease and flexibility of data set manipulation.....SPSS is the best**
- xii. Pricing for individualsR is free**
- xiii. Thoroughness of documentation....Stata is the best**

Popularity of Software based on number of scholarly article in 2016



SPSS

Data Format and Compatibility

- .sav file to save data
- .spo to save output files
- Syntax files (.sps) for programming
- Import
 - Excel files (.xls, .xlsx),
 - Text files (.csv, .txt, .dat),
 - SAS (.sas7bdat),
 - Stata (.dta)
- Export
 - Excel files (.xls, .xlsx),
 - Text files (.csv, .dat),
 - SAS (.sas7bdat),
 - Stata (.dta)

SPSS

- Strengths

- Easy and intuitive user interface; menus and dialog boxes
- Similar feel to Excel
- SEMs through SPSS Amos
- Easily exclude data and handle missing data

- Limitations

- Absence of certain robust methods (e.g...Least Absolute Deviation Regression, Quantile Regression, ...)
- Some of the more complex procedures are only available through the use of syntax, a programming language

Stata

Data Format and Compatibility

- .dta file to save dataset
- .do syntax file, where commands can be written and saved
- Import
 - i. Excel files (.xls, .xlsx),
 - ii. Text files (.txt, .csv, .dat),
 - iii. SAS (.XPT),
 - iv. Other (.XML)
- Export
 - i. Excel files (.xls, .xlsx),
 - ii. Text files (.txt, .csv, .dat),
 - iii. SAS (.XPT),
 - iv. Other (.XML)
- Old versions of Stata cannot read newer versions of Stata datasets

Stata

Strength

- Syntax mainly used, but menus are an option as well
- Some user written programs are available to install
- Works well with panel data, survey data, multiple imputations etc...
- Data management

Limitations

- Moderately steep learning curve
- Can only hold one dataset in memory at a time
- Cannot handle very large datasets - may have to sacrifice the number of variables for the number of observations
- Graphs have limited flexibility

SAS

Data Format and Compatibility

- Available for Windows only
- Import
 - i. Excel files (.xls, .xlsx),
 - ii. Text files (.txt, .dat, .csv),
 - iii. SPSS (.sav),
 - iv. Stata (.dta),
 - v. JMP (.jmp),
 - vi. Other (.xml)

- Export
 - i. Excel files (.xls, .xlsx),
 - ii. Text files (.txt, .dat, .csv),
 - iii. SPSS (.sav),
 - iv. Stata (.dta),
 - v. JMP (.jmp),
 - vi. Other (.xml)

SAS

Strength

- BASE SAS contains the data management facility, programming language, data analysis and reporting tools
- SAS Libraries collect the SAS datasets you create
- More than 200 components are available to complement Base SAS which include SAS/GRAPH, SAS/PH (Clinical Trial Analysis), SAS/ETS (Econometrics and Time Series), SAS/Insight (Data Mining) etc...
- Handles extremely large datasets
- With one procedure, test results, post estimation and plots can be produced
- Size of datasets analyzed is only limited by the machine

Limitations

- Steep learning curve
- Graphics can be cumbersome to manipulate
- Documentation and books tend to be very technical and not necessarily new user friendly

R

Data Format and Compatibility

- Import

- Excel files (.xls, .xlsx),
- Text files (.txt, .dat, .csv),
- SPSS (.sav),
- Stata (.dta),
- SAS(.sas7bdat),
- Other (.xml, .json)

- Export

- i. Excel files (.xlsx),
- ii. Text files (.txt, .csv),
- iii. SPSS (.sav),
- iv. Stata (.dta),
- v. Other (.json)

R

Highlights

- R is free and open source
- Over 6000 user contributed packages available through [CRAN](#)
- Large online community
- Network Analysis, Text Analysis, and Data Mining
- Interacts with other software such as, Python, Bioconductor, WinBUGS, JAGS etc...
- Scope of functions: flexible, versatile
- Size of datasets analyzed is only limited by the machine

Limitations

- Steep learning curve
- Large online help community but no 'formal' tech support
- Have to have a good understanding of different data types before real ease of use begins
- Many user written packages may be hard to sift through

Top free Statistics Software packages

- **R**
- **Epi Info**: A completely free statistical software for epidemiology developed by Centers for Disease Control and Prevention (CDC). It has a GUI user interface
- **PSPP**: free alternative to SPSS.
- **GNU Octave**: This tool presents an excellent alternative to Matlab.
- **Regressit**: A completely free add-in for Excel, Regressit can be used for multivariate descriptive data analysis and multiple linear regression analysis.
- **JASP**: free and open-source graphical program for statistical analysis, designed to be easy to use, and familiar to users of SPSS. Additionally, JASP provides many Bayesian statistical methods.
- **WinPEPI**: Can be downloaded free from <http://www.brixton-health.com>. It is different from most other statistic programs because it is not a spreadsheet. It only takes tabular data. For example, you can input the figures in a crosstab table into WinPEPI and it will give you not only the chi-square, but the Odds Ratio and the Relative Risk as well

Please see <http://statpages.info/javasta2.html> for a robust list of free statistics software

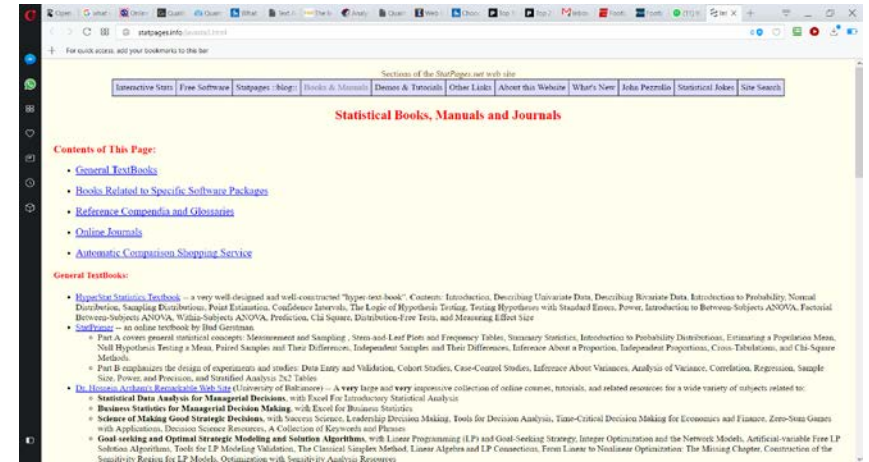
Online (web based) Software Packages

- [OpenEpi Version 2.2](#): Completely free
- [SOCR -- Statistics Online Computational Resource](#): Comprehensive and free
- [SciStatCalc](#): Also free
- [MedCalc.org](#): Also free
- [SISA \(Simple Interactive Statistical Analysis\)](#): Free. SISA allows you to do statistical analysis directly on the Internet
- [Statgraphics Stratus](#): Free version is limited

Please see <http://statpages.info/index.html> for a robust list of web based software

Useful Statistic links

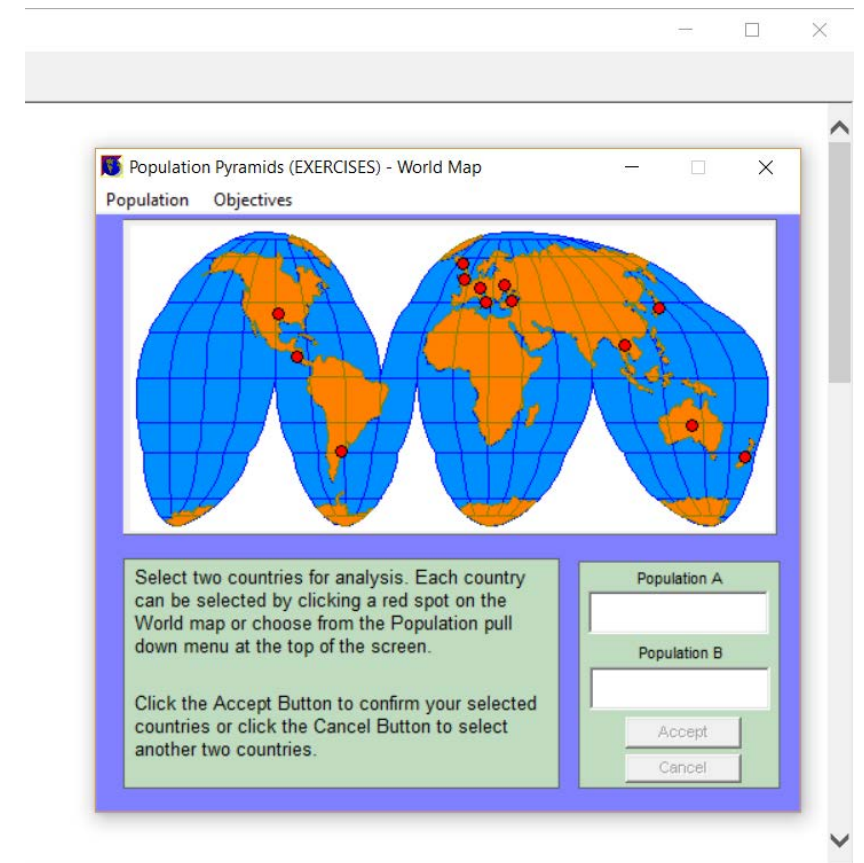
- <http://statpages.info/javasta3.html>: A link to very useful Statistical Books, Manuals and Journals
- <http://statpages.info/javasta4.html>: A link to very useful Interactive Statistical Demonstrations and Tutorials
- Don't know which statistical method to use? This interactive webpage may be helpful:
<http://www.microsiris.com/Statistical%20Decision%20Tree/default.htm>



- <http://statpages.info/javasta3.html>

Learning Epidemiology (and Statistics): The Fun way

- **Studying Populations:** This is a wonderful program for those intending to learn the basic principles of epidemiology without too much hassles.
- Download it from <http://medicine.dundee.ac.uk/studying-populations-epilex> and have a fun-filled adventure in learning statistics



Finally: A book that will definitely help....

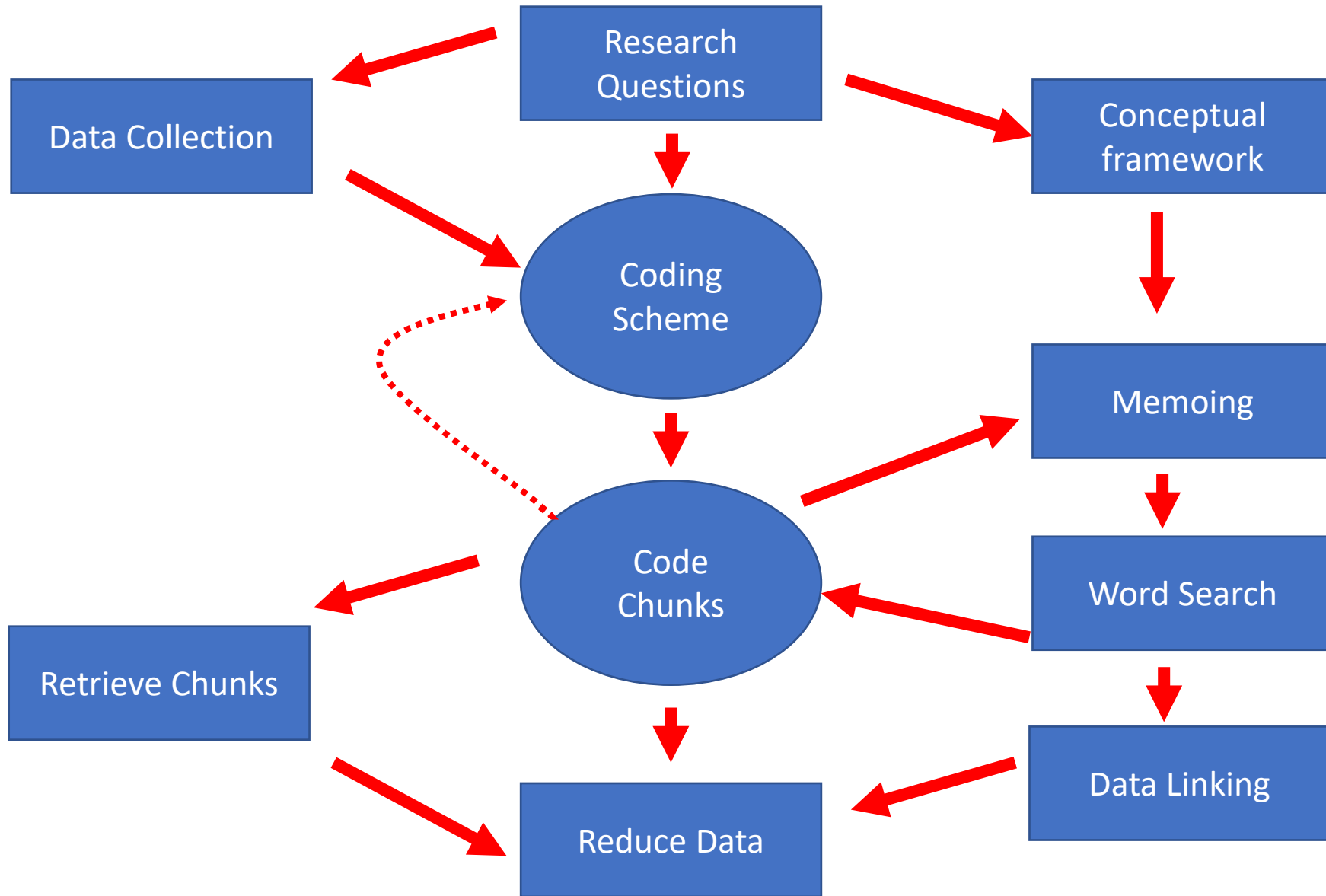
- Probably, the best selling book on the subject in Nigeria
- Currently in its fourth edition
- Covers basic and intermediate procedures in SPSS
- Includes supplementary chapters on:
 - Literature searches with HINARI, PubMed and Google Scholar
 - EndNote, the reference manager
- And lots of other goodies!

2. Qualitative Analysis Software

CAQDAS: Computer Assisted Qualitative Data Analysis Software



What are the steps involved in qualitative data analysis



Many of these tasks can be greatly assisted by computer software

Here is how qualitative software can help:

1. Making notes in the field;
2. Writing up or transcribing field notes;
3. Editing: correcting, extending, or revising field notes;
4. Coding: attaching key words or tags to segments of text to permit later retrieval;
5. Storage: keeping text in an organized database;
6. Search and retrieval: locating relevant segments of text and making them available for inspection;
7. Data "linking": connecting relevant data segments to each other to form categories, clusters, or networks of information;

8. Memoing: writing reflective commentaries on some aspect of the data as a basis for deeper analysis;
9. Content analysis: counting frequencies, sequences, or locations of words and phrases;
10. Data display: placing selected or reduced data in a condensed, organized format
11. Conclusion-drawing and verification: aiding the analyst to interpret displayed data and to test or confirm findings;
12. Theory-building: developing systematic and coherent explanations of findings and testing hypotheses;
13. Graphic mapping: creating diagrams that depict findings or theories.
14. Report-writing

Broad classification

A complete CAQDAS should encompass the following:

- Data Capture
- Text Retrievers
- Textbase Managers
- Code-and-Retrieve Programs
- Code-based Theory Builders

Note that some software focus on just one of these functions while some are designed as one-stop solution to qualitative analysis

Data Captors



Entering data

While there are specialized software for each of these data types (e.g. MS Word for Typing), most modern qualitative software integrate all these methods.

Dedicated transcribing software

AnnoTape 2.0

- This is a free software for recording, analysing, and transcribing audio, video, image and text data for qualitative research.
- AnnoTape turns your computer into a virtual tape or video recorder:
 - Record video or sound files - interviews, conversations, broadcasts - direct to your hard disk
 - Store up to one hundred hours of video, image, sound, together with text-based data, all in one integrated database
 - Analyse data by annotating and indexing the original media files
 - Effectively creating snippets of virtual tape for a database of data
 - At the touch of a button, return to the exact moment in the recorded data you seek

Please see <http://textanalysis.info/pages/text-analysis-software---classified/transcribing-software.php> for more examples

Text retrievers

- Find all instances of words and phrases in one or several files.
- Search for places where two or more words or phrases coincide within a specified distance (a number of words, sentences, pages, etc.) and allow you to sort the resulting passages into different output files and reports.
- They may do other things as well, such as content analysis functions like counting, displaying keywords in context, or creating.
- While there are dedicated text retrievers, function is present in most popular qualitative data software.

Textbase Managers

- Helps in organizing your data
- Good for annotating data for coding and memoing
- Also have good search capabilities

Code-and-Retrieve Programs

- These programs allow the researcher to apply category tags (codes) to passages of text, and later retrieve and display the text according to the researcher's coding.
- Has completely replaced the traditional scissors and paper approach to coding and memoing

Code-based Theory Builders

- These are higher order code and retrieve programs
- They have more powerful memoing and coding features that allow researchers to formulate and test theories from data.
- Examples include AFTER, AQUAD, ATLAS/ti, Code-A-Text, HyperRESEARCH, NUD*IST, QCA, The Ethnograph, and winMAX

Free CAQDAS

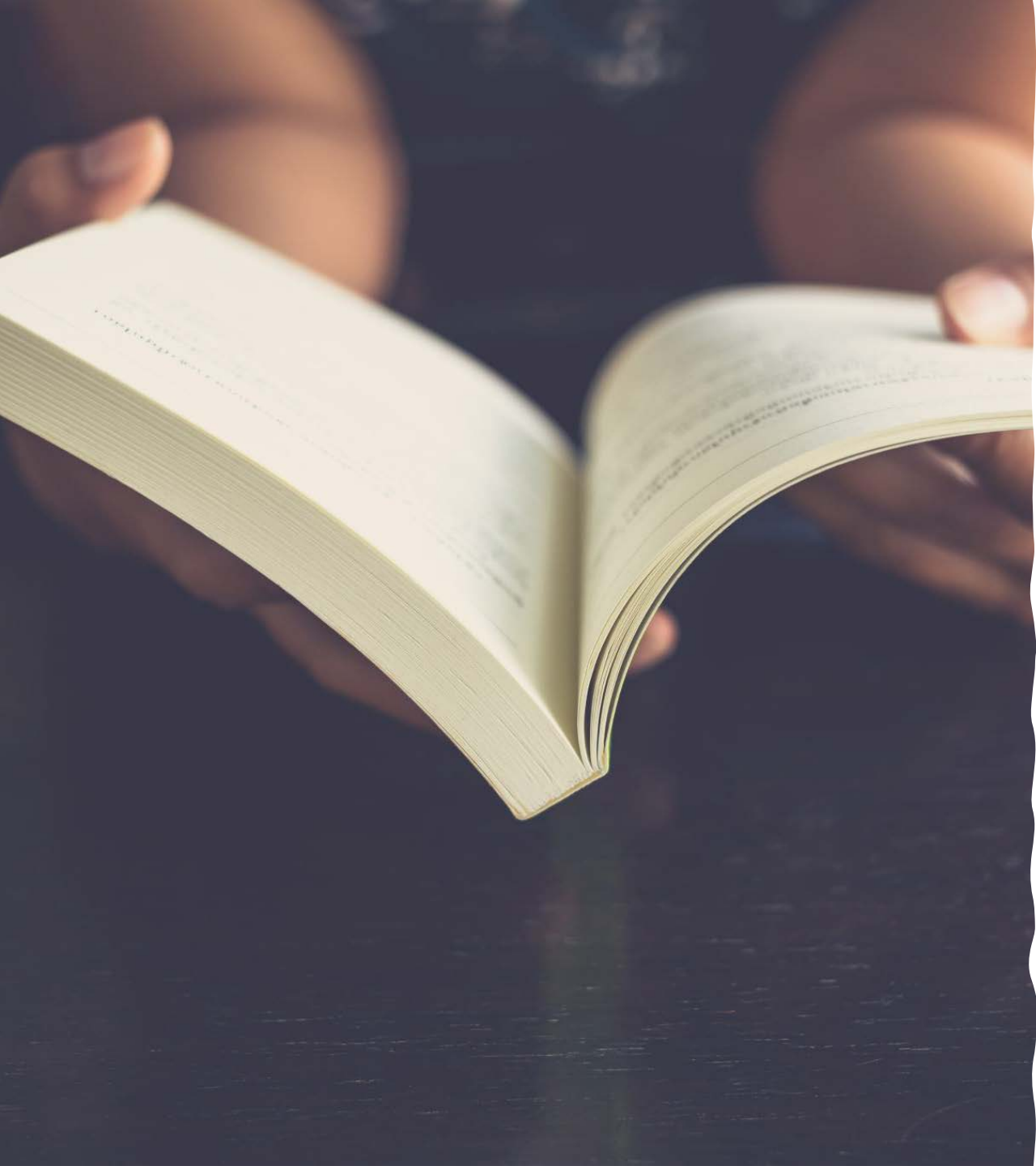
- **Aquad** (<http://www.aquad.de/en/>)
- **Coding Analysis Toolkit (CAT)**: CAT was the 2008 winner of the "Best Research Software" award from the organized section on Information Technology & Politics in the American Political Science Association. (<http://cat.texifter.com>)
- **RQDA**: This is the qualitative plugin to R, the statistic software. What a plugin to such a power statistics program means is that implementing quantitative data to your project will be seamless and effective (<http://rqda.r-forge.r-project.org>)
- **QDA Miner Lite**: A "lite" version of the proprietary software, QDA Miner Lite has a slightly limited functionality
- **Open Code**: is a tool for coding qualitative data generated from text information such as interviews, observations or field notes. Open Code was originally developed to follow the first steps of the Grounded Theory methodology but this version can as well be used for Qualitative Content Analysis. <http://www.phmed.umu.se/enheter/epidemiologi/forskning/open-code/>

Proprietary CAQDAS

- **ATLAS.ti**: Features include intuitive interface, data management, coding, memos and comments, hyperlinking, visualization, advanced concept searching, survey import, optimized for teamwork.
- **MAXQDA**: Features include text analysis, data management, import of variables and survey data, coding, memos, mixed-methods analysis, optimized for teamwork
- **NVivo 10**: Features include text analysis, multilingual interface, interface is designed similarly to Microsoft Office Suite, advanced text searching, multiple output formats, import web pages and social media datasets.
- **QDA Miner**: Features include text management, coding, text retrieval and coding analysis, multilingual interface, optimized for teamwork, easy to save output in variety of formats.

Some Controversy about CAQDAS

- Disengagement from the data is a potential concern in using software packages (Banner & Alberran, 2009; St.John & Johnson, 2000).
- Researchers may focus on the process of the technique instead of the meaning of the data.
- Transcribing interviews gives the researcher the opportunity to become immersed in the richness of the data, whereas using the CAQDAS can become cumbersome with the number of codes and categorizations created by the software



Final thoughts.....

When reading a journal article.....

Keep In Mind That

- No study is perfect
- All data is dirty in some way or another; research is what you do with that dirty data
- Measurement involves making choices

Be Critical About Numbers

- Every statistic is a way of summarizing complex information into relatively simple numbers.
- How did the researchers arrive at these numbers?
- Who produced the numbers and what is their bias?
- How were key terms be defined & in how many different ways?

Be very critical about numbers

- How was the choice for the measurement made?
- What type of sample was gathered & how does that affect result?
- Is the statistical result interpreted correctly?
- If comparisons are made, were they appropriate?
- Are there competing statistics?

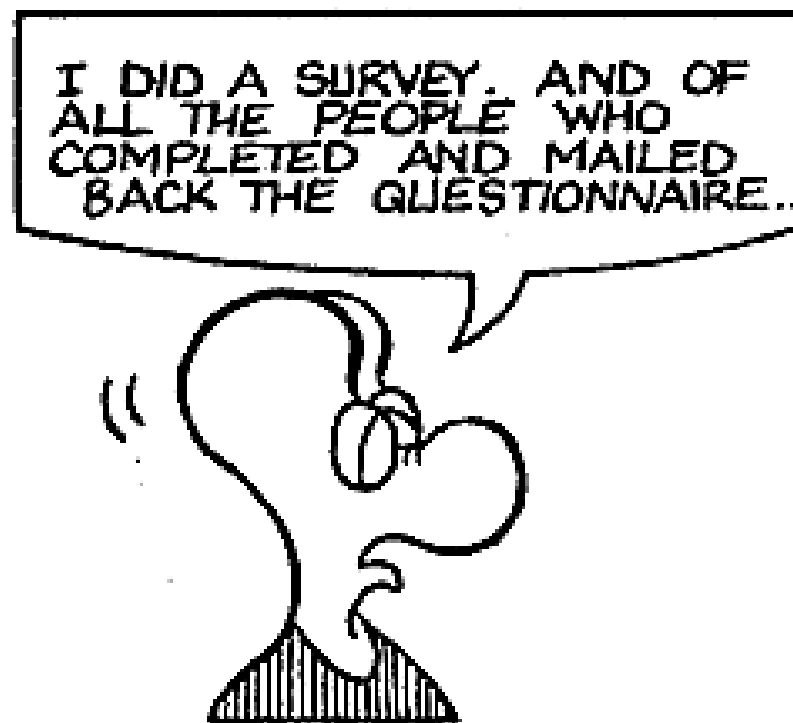
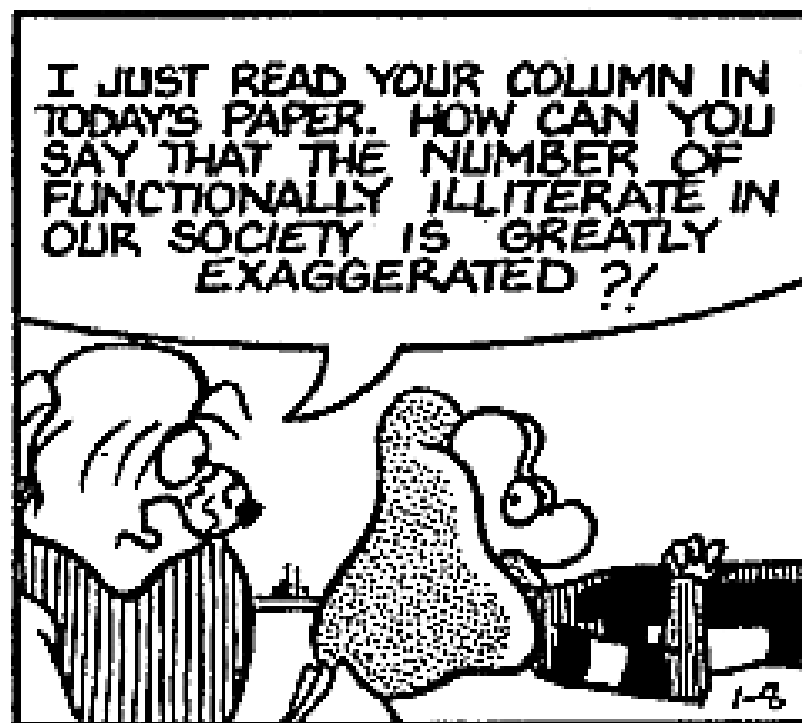
Be really critical About Numbers

- With one foot in a bucket of ice water, and one foot in a bucket of boiling water, you are, on the average, comfortable.



Be critical about numbers: Bias and Error

THE MICE SQUAD





Thanks for your attention

To ask questions, please join the
forum at www.oluwadiya.com