LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Violeta Carrión, M.Ed, MA

Data Analyst and Statistical Research Programmer
LinkedIn: violetacarrion * violetacarrion.com
Facebook Groups:
/statausersgroup * /raschmeasurement */ systatusersgroup
vdcarrion@gmail.com

BIOSTATISTICS II
CRME 421
Time: 5:00-8:00 pm
Location: CTRE. room 504
Ramón A. Durazo-Arvizu, PhD.
Professor
Public Health Sciences
(708) 327-9007
rdurazo@luc.edu

1/ 54

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Outline

- Humanware Introductions
- Statistical Tools of the Trade: or the Tale of the many S's.
- Walkthrough of Stata's Windowing system
- Overview of Stata command line programming

2/ 54

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Outline

- Statistical Tools of the Trade: or the Tale of the many S's.

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

1. Statistical Analysis Software

- The development of SAS began in **1966** after North Carolina State University re-hired Anthony Barr to program his analysis of variance and regression software so that it would run on IBM System/360 computers. The project was funded by the National Institute of Health and was originally intended to analyze agricultural data to improve crop yields. The first versions of SAS were named after the year in which they were released. In 1971, SAS 71 was published as a limited release. It was used only on IBM mainframes and had the main elements of SAS programming, such as the DATA step and the most common procedures in the PROC step. In 1976, Barr, Goodnight, Sall, and Helwig removed the project from North Carolina State and incorporated it into SAS Institute, Inc.

https://en.wikipedia.org/wiki/SAS_(software):

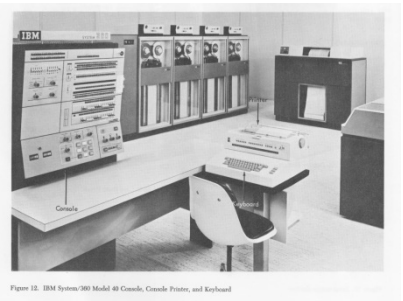


Figure 12. IBM System/360 Model 40 Console, Console Printer, and Keyboard

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

2. Statistical Package for the Social Sciences

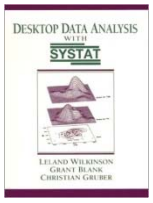
- The software was released in its first version in 1968 as the Statistical Package for the Social Sciences (SPSS) after being developed by Norman H. Nie, Dale H. Bent, and C. Hadlai Hull. Those principals incorporated as SPSS Inc. in 1975. Early versions of SPSS Statistics were designed for batch processing on mainframes, including for example IBM and ICL versions, originally using punched cards for input. A processing run read a command file of SPSS commands and either a raw input file of fixed format data with a single record type. From version 10 (SPSS-X) in 1983, data files could contain multiple record types.
- http://www.ibm.com/analytics/us/en/technology/spss/spss-trials.html

SPSS AN IBM COMPANY

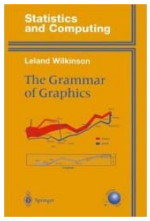
5/ 54

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

3.The SYStem for STATistics



SYSTAT is a statistics and statistical graphics software package, developed by Leland Wilkinson in the late 1970s, who was at the time an assistant professor of psychology at the University of Illinois at Chicago. Systat was incorporated in 1983 and grew to over 50 employees.
In 1995 SYSTAT was sold to SPSS Inc., who marketed the product to a scientific audience under the SPSS Science division. By 2002, SPSS had changed its focus to business analytics and decided to sell SYSTAT to Cranes Software in Bangalore, India. Cranes formed Systat Software, Inc. to market and distribute SYSTAT in the US, and a number of other divisions for global distribution. The headquarters are in Chicago, Illinois.
By 2005, SYSTAT was in its eleventh version having a revamped codebase completely changed from Fortran into C++. Version 13 came out in 2009, with improvements in the user interface and several new features.



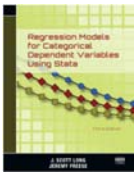
6/ 54

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

4. STATA

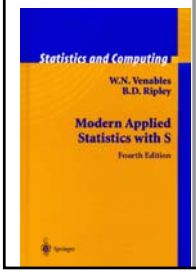
Stata is a general-purpose statistical software package created in 1985 by StataCorp. Most of its users work in research, especially in the fields of economics, sociology, political, science, biomedicine and epidemiology. Stata's capabilities include data management, statistical analysis, graphics, simulations, regression, and custom programming.
The name *Stata* is a syllabic abbreviation of the words *statistics* and *data*.[2] The FAQ for the official forum of Stata insists that the correct English pronunciation of *Stata* "must remain a mystery".

- There are four major builds of each version of Stata:[4]
  - Stata/MP for multiprocessor computers (including dual-core and multicore processors)
  - Stata/SE for large databases
  - Stata/IC, which is the standard version
  - Small Stata, which is a smaller, student version for educational purchase only

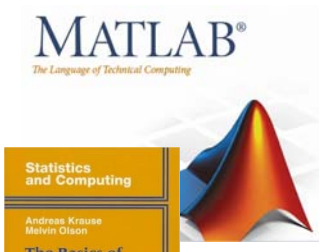


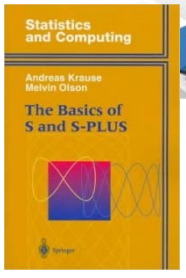
LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Other titles…



- **https://en.wikipedia.org/wiki/List_of_statistical_packages**
- **https://stat.utexas.edu/training/software-information**
- **Journal of Statistical Software**
  - https://www.jstatsoft.org/index





LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Resources for Stata

- http://www.ats.ucla.edu/stat/stata/
- http://www.stata.com/links/
  - Includes videos
- http://www.cpc.unc.edu/research/tools/data_analysis/statatutorial
- http://www.ssc.wisc.edu/sscc/pubs/sfr-intro.htm
- http://www.statalist.org/
- http://data.princeton.edu/stata/default.html
- http://statadaily.com/
- https://stat.utexas.edu/training/software-tutorials#stata
- https://www.facebook.com/groups/statausersgroup/

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Things to remember ...

- Missing values = infinite positive
  - flag > 5 will include missing
  - flag > 5 & !missing(flag) will not
- Case sensitive
- Web aware
- Stores data ALL in memory
- Works much like an Unix system

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Outline

- Walkthrough of Stata's Windowing system

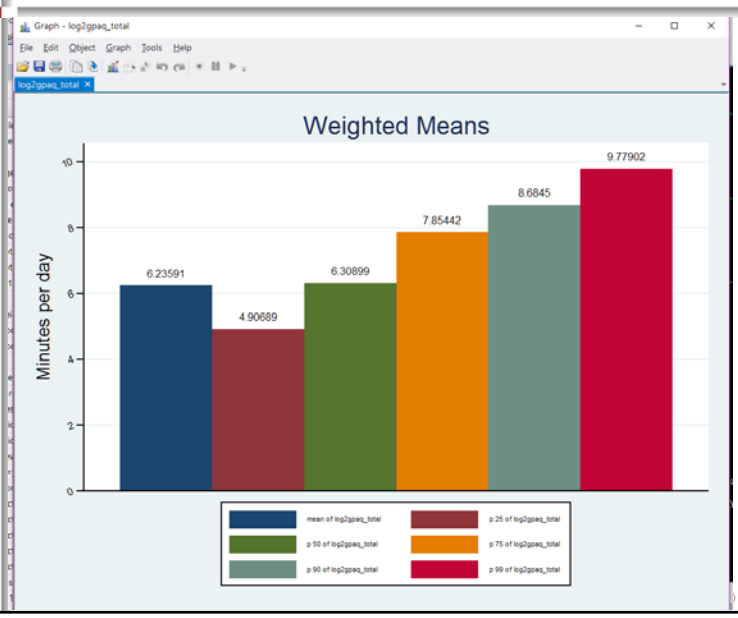
LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Walkthrough of Stata's Windowing system

- Command
- Results
- Review
- Variables
- Properties
- Graph
- Viewer
- Data Editor
- Do file Editor
- Variables Manager

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Graph and Graph Editor Window



13/

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Outline

- Walkthrough of Stata's command line and programming
  - DO files

14/

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

Commands to know…

- findit or net search or help –find information or programs
- ls or dir – List files in directory
- cd -- change directory location
- list --List value of variables
- summarize --Summary statistics
- describe --Describe data in memory or in file codebook Describe data contents
- tabulate --Tabulate frequencies
- generate --Create or change contents of variables
- egen Extensions to generate
- correlate --Correlations (covariances) of variables or coefficients
- ttest --Mean-comparison tests
- regress --Linear regression
- alpha --Compute interitem correlations (covariances) and Cronbach's alpha
- graph --The graph command

15/

LOYOLA MEDICINE
Loyola University Chicago Stritch School of Medicine

16/