Psychology 350: An introduction to R for Psychological Research
History of computing

William Revelle
Northwestern University
Evanston, Illinois USA

https://personality-project.org/courses/350

Spring, 2024

Outline

Part I: What is R, where did it come from, why use it

- Installing R and adding packages: the building blocks of R

Part II: A brief introduction – an overview

- R is just a fancy (very fancy) calculator
- Descriptive data analysis
- Some inferential analysis

Part III R is a powerful statistical system

- Data entry (detail and practice)
- Descriptive (again)
- Inferential (t and F with more practice)
- Regression
- Basic R commands

Part IV: Psychometrics

- Reliability and its discontents
- EFA, CFA, SEM

Part V: Help and More Help

- List of useful commands

Part VI: The psych package and more practice

Ada Lovelace: The first programmer

1. The Babbage Analytical Engine (1838) was a design for a punch card based calculating machine (generalized from an automatic loom) but was never actually built. (see Bromley, 1982).
2. Ada Lovelace wrote a "memoir" on how to program it (Lovelace, 1842). She was perhaps the first programmer.
3. The U.S. Department of Defense (1980) initiated an operating system/language in her honor (ADA) to get around the problem of too many different programming systems.

The von Neuman machine and programable analysis

1. The "Von Neuman" machine instigated programs that could act depending on the state of the machine. (See Isaacson, 2014)
   - Developed to calculate flight paths for artillery and rockets.
   - Turing had developed computers for breaking codes, but this work was classified and not known about until much later.
2. Large scale computing was for rocket design, military applications, the census, meterology.
3. "Computers wore skirts" at NASA (Johnson, nd: see Katherine Johnson's 90th birthday Computers were operating desk calculators.
4. The "computers" could do in a morning what the engineers would take all day to do NASA history 1935-1970

Computers go to college/university

1. The use of computers in physics, oceanography, and psychology exploded in the 1950s-1960s, as it did in business
2. Originally using electronic tubes
   - Univac (1950), IBM 704 (1954), 7090 (1959)
3. Control Data Corporation and the "super computer"
   - CDC1604 (designed by Seymour Cray) was the first commercially successful computer with transistors (1960)
   - CDC 3600 (1963)
   - CDC 6400 and 6600 (1964)
4. IBM 360 series (vapor ware?)
5. These were all large systems that eventually allowed remote time-sharing.
6. Commands were submitted using punched cards (Hollerith aka IBM cards)
7. Turn around time ranged from minutes (late at night) to 8-12 hours (during the day).

Computer languages (a brief summary

1. Machine language (written in Boolean at the bit level), organized as a very limited set of instructions
2. Assembly languages would take somewhat higher level instructions and generate the machine language translation
3. Compiled languages (e.g. FORTRAN, LISP, C) were more readable, but needed to be "compiled" into Assembly which then in turn would run machine level commands.
4. Operating systems (OS) provide higher level functionality, but tend to be written in Assembly.
5. Large sets of programs would be organized for specific tasks; e.g., for statistics, forecasting, modeling.

Statisical computing languages for mainframes

1. BMD and BMDP for biomedical research (UCLA)
2. SAS for agricultural research (North Carolina State)
3. SPSS (Statistical Package for the Social Sciences) (Stanford)
   - SPSS was written in Fortran for the IBM family of computers
   - Northwestern University was a major source of translation of SPSS into CDC Fortran
4. Although originally developed at universities, these three statistical systems eventually became stand alone and proprietary companies.

From mainframes to minis to micros to personal computers

1. The three main statistical systems were written in Fortran for mainframe computers.
2. Even with the introduction of remote terminals, the main frame syntax remains in these languages.
3. With the introduction of personal computers, statistics could now be done interactively.
4. Languages for personal computers included SysStat (by Leland Wilkinson) as well as S (developed at Bell Labs) and S+.
5. R owes its history to S (originally conceived as S for the Mac).

And now, to R

A brief history of Computers and Psychology
R: Intro
R: Intro part 2

Bromley, A. G. (1982). Charles Babbage's Analytical Engine, 1838. *IEEE annals of the history of computing*, *4*(3), 196–217.

Isaacson, W. (2014). *The Innovators: How a Group of Inventors, Hackers, Geniuses and Geeks Created the Digital Revolution*. Simon and Schuster.

Lovelace, A. A. (1842). Sketch of the analytical engine invented by Charles Babbage, by LF Menabrea, officer of the military engineers, with notes upon the memoir by the translator. *Taylor's Scientific Memoirs*, *3*, 666–731.