



# DATA ANALYSIS IN R

“let the data speak for themselves” -- Tukey

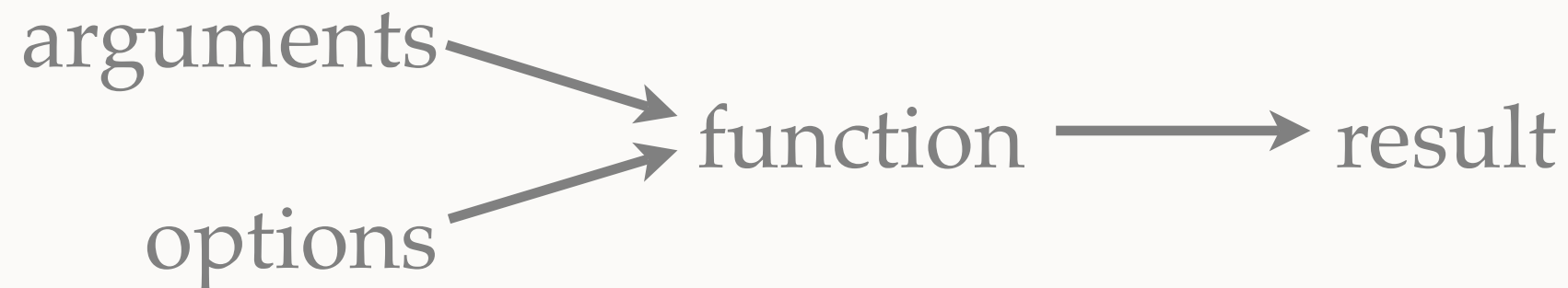


# WHAT IS R?

- free software system (language) for statistical analysis and graphics that can run on Unix, Windows, Mac OS
- developed and supported by a world-wide group of users
- various components (packages) available to extend functions
- able to import/ export data and analysis in many formats
- because R is a computer language, the process of analysis can be easily recorded as a 'script' that can be rerun with new data and easily shared with and edited by others

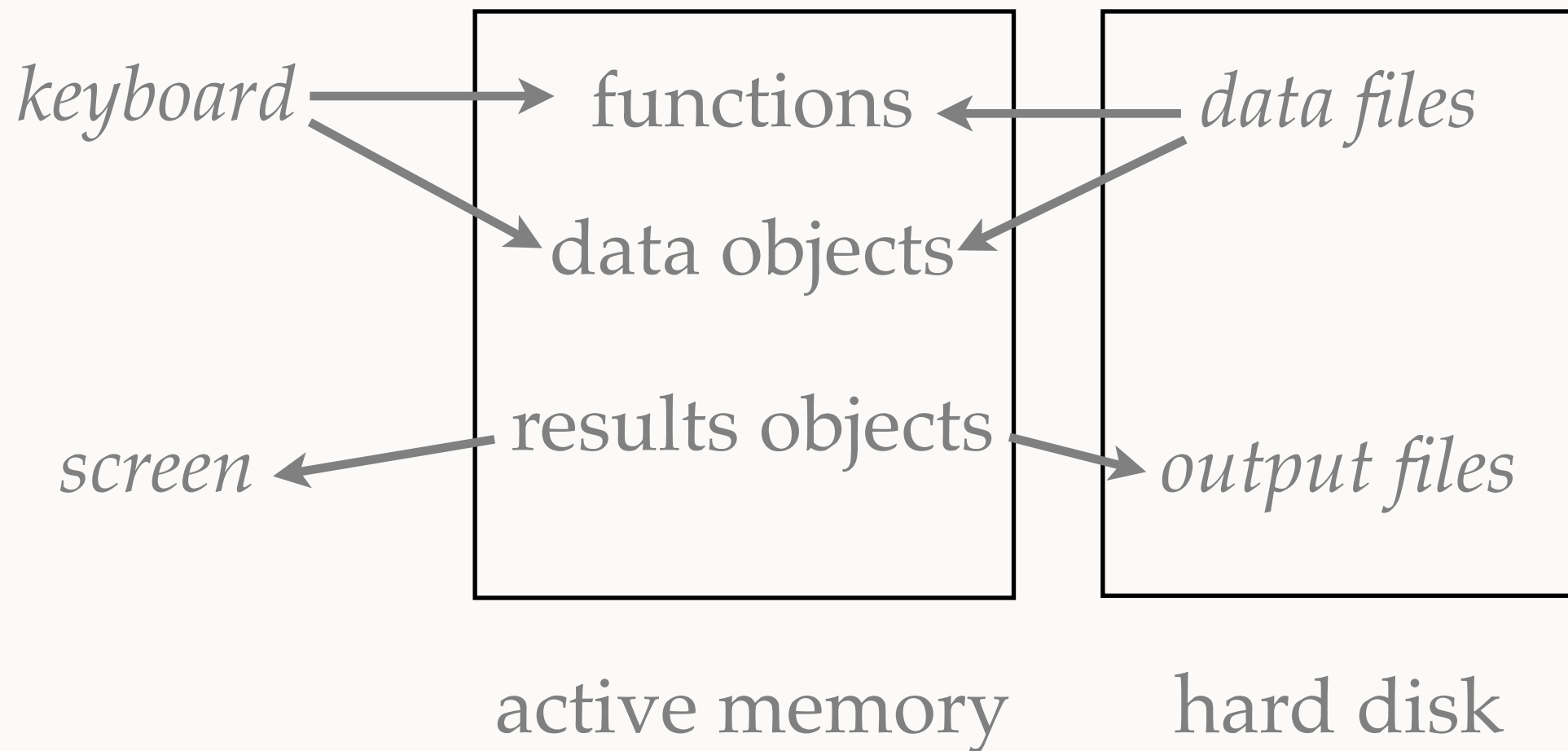
# HOW R WORKS

- R is interpreted, not compiled
- the syntax is simple (really!) and based on running pre-installed or user-defined functions



- `>function(arg, opt)`

# HOW R WORKS



■ `>ls()`



# ATTRIBUTES AND TYPES OF OBJECTS

- all objects have two intrinsic attributes: *mode* and *length*
- the four main *modes* are: numeric, character, complex, logical
- there are several object types that can be used to represent data: vector, factor, array, matrix, data frame, time series and list

# WORKING WITH OBJECTS

- create objects
  - assign operator: <-
- display information about objects
  - ls(); ls.str(); mode(); length(); attributes(); dim()

# IMPORTING DATA INTO OBJECTS

- `read.table()`; `read.csv()`
- Excel files (and all other spreadsheets, databases, etc) can save data as simple text (e.g. csv) files
  - some pitfalls -- spaces in column names, non text characters
- Editing in memory: `edit()`
- practice



# WORKING WITH DATA IN OBJECTS

- use '\$' to refer to specific variable in data frame (from `names(data.frame)`)
- practice
- example: from Wanich et al.



# REMOVING OBJECTS/ SAVING WORK

- `rm(); rm(list)`
- saving an analysis file
- saving a work space

# ONLINE HELP

- `?; help() -- try.all.packages` option
- package manager; package installer



# REFERENCES

- Clarkson University
  - <http://www.cyclismo.org/tutorial/R/>
- Emmanuel Paradis