Comparison of R and SAS

R vs. SAS

R

- Free
- No warranty
- Open source (CRAN) - 9062 packages and growing @ 2000/year
- No GUI
- Function/object based programming
- Nice, simple graphics

SAS

- Annual license fee
- Excellent technical support
- Corporate R&D/QC
- Enterprise Guide GUI
- Conversational programming
- Complex graphics – GUI available

R vs. SAS

R

- Loops are slow -> Vectorize!!!
- Limitations on the types of data that R handles well
- Unable to handle large data sets quickly and efficiently
- No method for writing out tabular reports
- Popular in Academia

SAS

- All code loops
- Robust interface to all commercial data
- Scalable/efficient for large data
- ODS writes reports to all common formats
- Used extensively in FDA drug trials

The Basics of R

The R Environment

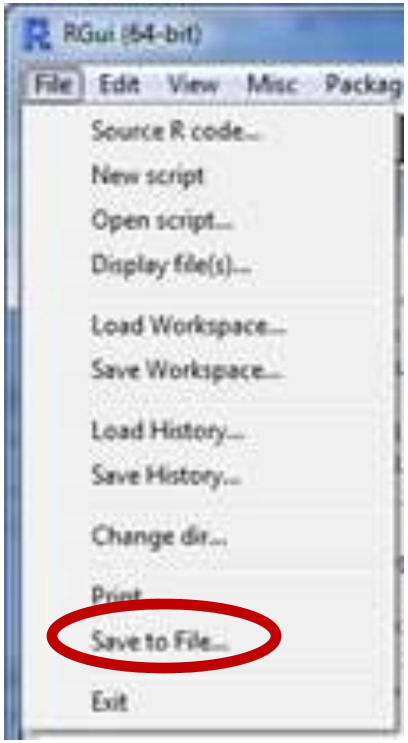
➢Console

- Accepts commands, displays results and messages
- > Command prompt
- + Incomplete command
- Case sensitive
- Use Up arrow to retrieve previous command
- Use ; or new line to separate commands
- Use # to designate comments

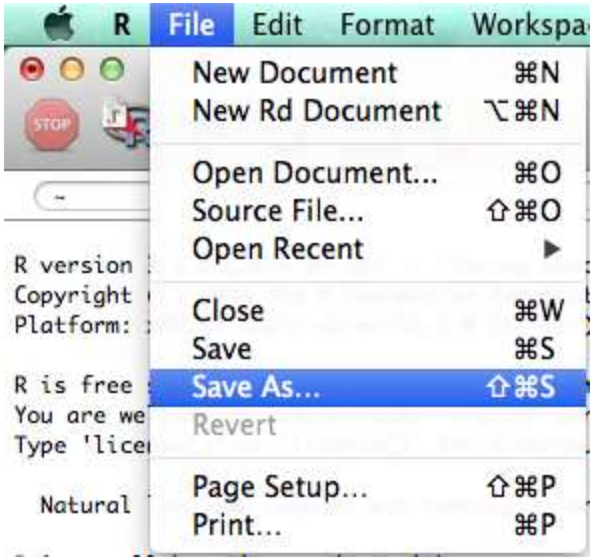
The R Environment

Saving the Console

- Windows



- Mac



```
R version
Copyright
Platform:

R is free
You are we
Type 'lice

  Natural
```

The R Environment

- Redirecting TEXT Output (results)
  - sink(“filename”, split=TRUE)
    - filename = name & path of output file
    - split=TRUE – sends output to both file and console
    - split=FALSE – sends output to file only
  - sink() - closes file

The R Environment

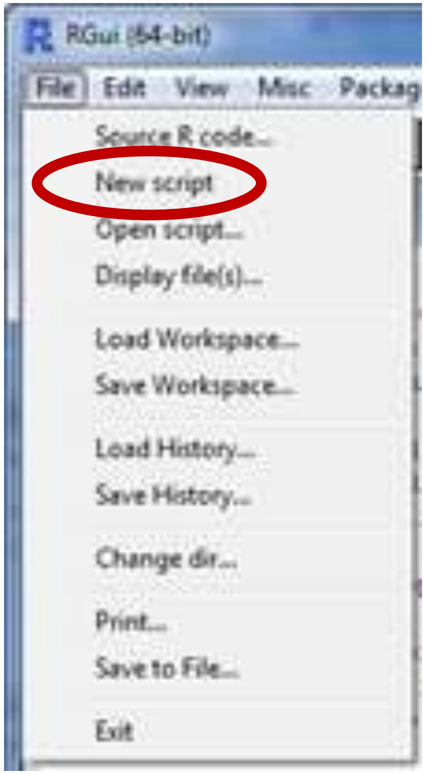
➢ Scripts

- Save and retrieve your work (R “programs”)
- Text files with .R extension
- Use R Editor, Word, Emacs, or other text editor
- Beware of Office quote marks & capitalization
- Use { } to extend a statement on multiple lines
- Ctrl+r (Windows) or command+return (Mac) executes highlighted script commands

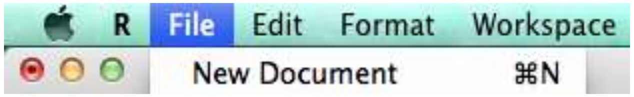
The R Environment

Opening the R Editor to Create a Script

- Windows



- Mac



The R Environment

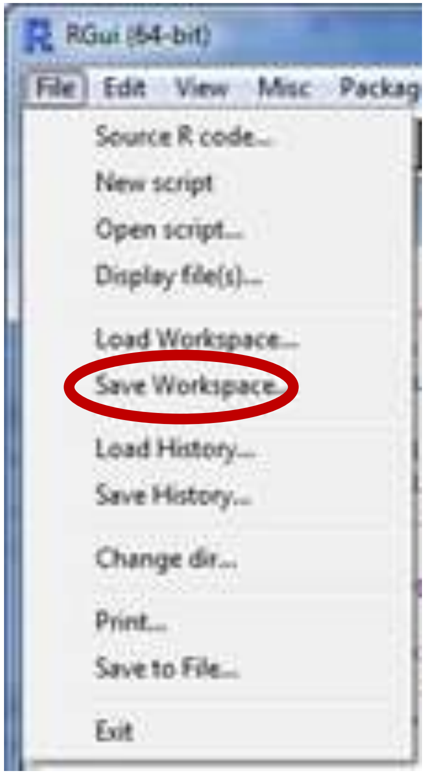
➢ The R Workspace

- Collection of objects currently stored in R
- Objects created or loaded during R session
- objects() or ls() – display contents of workspace
- May be saved as file – has RData extension
- Saving when prompted at end of session creates both .Rdata and .Rhistory files
- R reloads these files next time you run R

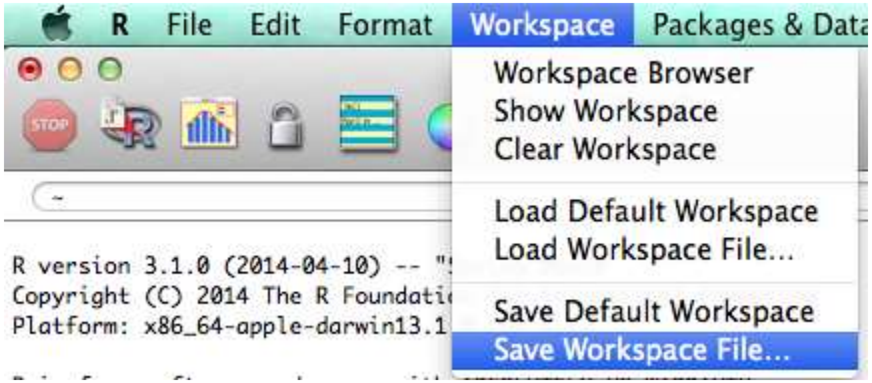
The R Environment

Saving the Workspace

- Windows



- Mac



```
R version 3.1.0 (2014-04-10) -- "
Copyright (C) 2014 The R Foundati
Platform: x86_64-apple-darwin13.1
```

The R Environment

Script Commands to Save and Load Workspaces

- Save

```
save.image("C:/Users/kinchelf/Documents/DemoWorkspace.RData")
```

Note: Windows can also use \\ in paths

getwd() – shows default path

- Load

```
load("C:/Users/kinchelf/Documents/Jan3.RData")
```

Good Housekeeping

- Know what you have - objects() or ls()
- Use meaningful names
- Clean up what you don’t need
  - rm(*UnwantedObject*) – removes single item
  - rm(list=ls()) – removes everything
    - – Use judiciously especially in shared environment
    - – Avoid if you have created functions you want to keep

The R Environment

- ➢Packages (libraries)
  - Extend functionality of R
  - May need to download and install package first
  - Use library() to show available packages
  - Use library(PackageName) to load
  - Use search() to see which packages are loaded
  - foreign – required for accessing external data such as SAS
  - boot – contains bootstrap functions

Getting Help

- ?*functionname* – opens help page for function
- *functionname* – displays code of function
- example(*functionname*) – gives examples
- demo(*functionname*) – demo of some functions
- ??*keyword* – opens possible help pages
- ??”*multiple words*”
- PDF documents
- The Internet

Data Storage in R

- Named data structures (objects)
- Vector – series of data values
- Scalar – single value vector
- Matrix or array – multidimensional vectors of same data type (matrix: 2 dimensions)
- Factor – grouping by category
- Data frame – matrix-like structure with different data types
- Function – object containing program code
- Typically returned by class() function

R Data Types

- Numeric
  - Integer
  - Double
- Logical – True/False
- Character
- List – elements not of same type (also a structure)
- Complex - (real + imaginary)
- Raw – data bytes represented as 2 hex digits
- Typically returned by mode() function

Information About Objects

- class(*objectname*) – reveals object structure
- mode(*objectname*) – reveals data type
- summary(*objectname*) – additional info depending on class of object
- str(*objectname*) – **str**ucture of R object
- length(*objectname*) – number of values

Quirky Things About R

- Many **R** objects have a class attribute, a character vector giving the names of the classes from which the object *inherits*. If the object does not have a class attribute, it has an implicit class, "matrix", "array" or the result of mode(x) (except that integer vectors have implicit class "integer").
- No class: class=mode