L05 R Programming III

EPID 799B

Fall 2016

Xiaojuan Li

Objective

- Packages
- Review basic programming in R
  - Syntax
  - Examine data frame
  - Extracting parts
  - Basic graphics

Announcements

- Homework 1 assigned and due Monday 9/19
- Short survey due Wednesday 9/14 https://goo.gl/forms/eCAg3oXefVqOQd4t2

Packages

Libraries or packages – how and where R organize statistical and data analyses

- **library() #** to see all packages available
- To install – click install packages in the Packages window and type foreign or type **install.packages("foreign")** in the command window.
- To load or activate – check box in front of the package (e.g foreign) or type **library("foreign")** in the command window

To run commands or a script

- To run a line, put your cursor on the line with the code, press Run or control+enter or command+enter (Mac)
- To run several lines, highlight some code, press Run or control+enter
- To run all codes, press Source with Echo

- Save your scripts although Rstudio stores then automatically
- Read results in the console
  - Results appear after an index number
  - Commands appear after the command prompt, >

Getting help

citation() # shows how to cite the R software in your paper comments

To get help in RStudio

- Click on the Help button, type keywords and search
- In console, type and run **?functionname** or **??functionname**
  - e.g. **?read.table**
  - **??read.table**
- **help.search("data input")**

Naming objects

A valid object name

- Must start with an alphabetical character
- May contain numeric characters thereafter
- May contain period
- Cannot have spaces
- Example
  - x.1
  - x1

“Housekeeping”

To remove all objects you have created

- **rm(list=ls(all = TRUE))**

To remove some objects you have created

- **rm(a, b, c)**

To clean the Console:

- ctrl + L
- **cat("\f")**
- **cat("\014")**

Group exercise

1. Read in the births dataset
2. Create a new data frame and name it births1
3. Examine the data frame births1
   a. Understand the data structure
   b. Print out the variable names
   c. Print out the first 10 observations
4. Create several frequency tables:
   martial marital*sex mrace*marital*sex
5. Plot the distribution of mage using a histogram
6. Subset the data to observations with 20<=mage<99 and sex not 9
7. Create a new variable called female: female=0 when sex=1; female=1 when sex=0; female is missing when sex=0
8. Make a frequency table of the new variable female