R Essentials

Christopher Manning

4 November 2007

Here are some collected pointers on some useful things that you should have picked up, but just to make sure that you have....

1 Data manipulation

1.1 Creating vectors

- `c(11, 19, 23, 7)` Creates a vector of the given items
- `seq(1, 10)` or `1:10` or `seq(1, 10, 0.1)` creates vectors that are sequences incrementing by 1 or the amount given
- `rep(c(3,6,9), 3)` and especially `rep(1:2, c(12, 19))` replicates a list k times, or each item in a list the number of times shown – useful for making things like group codes for later selection.
- `matrix(c(4, 8, 11, 2, 9, 7), nrow=3, byrow=T)` is one way to make a matrix.
- `rbind(c(4,8), c(11,2),c(9,7))` or `cbind(c(4, 11, 9), c(8, 2, 7))` make the same matrix by gluing rows or columns together.

1.2 Manipulating data frames

While you can get out and manipulate parts of a data frame using array indexing and conditional expressions, usually there are easier to use functions.

- `spdative <- subset(dative, dative$modality=="spoken")` The subset() function gives you a subset of a data frame. The first argument is the data set, the second aargument is an expression to select rows, and an optional third argument will select columns.
- `spdative <- transform(spdative, ThemeMinusRecipient = LengthOfTheme - LengthOfRecipient)` The transform() function lets you add columns to a data frame. (Think of it that you are adding columns with variable transformations, as here.) You can either save the result in the same data frame object or in a new object.

2 Plots

2.1 Saving plots

- You make a pdf plot like this (the filename is up to you; this is for Unix):

  ```
  pdf("~/ling289/example.pdf")
  plot(x, y) dev.off()
  ```

  It's a slightly awkward recipe, but it's what you do. You can do the same with many other file types. Other good ones to know are postscript(), xfig(), and jpeg().

2.2 Adding stuff to plots

- `lines(c(150000, 200000, 250000), c(59.5,60,61))` Add lines to a plot joining the points specified by two vectors
- `abline(h=60.25)`, `abline(v=200000)`, `abline(a=59.75, b=0.1)`, and `abline(lm(mt03.ter ~ train.vocab))` Add lines to a plot horizontally, vertically, with intercept and slope, and according to the best fit of a linear model