Bio Stats II : Lab 1

Gavin Fay

01/27/2016

Lab schedule

1/27: Introduction to R and R Studio, working with data
2/03: Intro to plotting, manipulating data
2/10: Probability, linear modeling, PCA
2/17: Programming practices, conditional statements
2/24: Creating functions, debugging
3/02: Permutation analysis
3/09: Advanced plotting

Why R?

Reproducible

- command line interface encourages organization
- scripts allow others (and you!) to reproduce analyses from end-to-end

Extensible

- new methods delivered as developed
- continual expansion through new packages

Open-source

- all code can be examined by the user

Free

- available to large set of users (and therefore developers)

R is not the only solution out there.
The real goal is not to teach R, but concepts that all programming depends on.

Trevor Branch rule:

"Every analysis you do on a dataset will have to be redone 10–15 times before publication.

Plan accordingly."

Recommended reading

An introduction to R (Venables et al.)
– `http://cran.r-project.org/doc/manuals/R-intro.pdf`

R reference card 2.0 (Baggott)
– `http://cran.r-project.org/doc/contrib/Baggott-refcard-v2.pdf`
– Extremely useful handout: put on wall in view of your desk

There are many (many) R books out there. Good for reference.

Installing R and R Studio

R (`http://r-project.org`)
- download appropriate version for your OS

R Studio (`http://rstudio.com`)
- a very good Integrated Development Environment (IDE) for R
provides:
- text editor
- syntax highlighting
- seamless code execution with R

You can use other text editors with R, but RStudio well organized.
(also looks same regardless of Operating System)

Getting started, command line

Getting Help

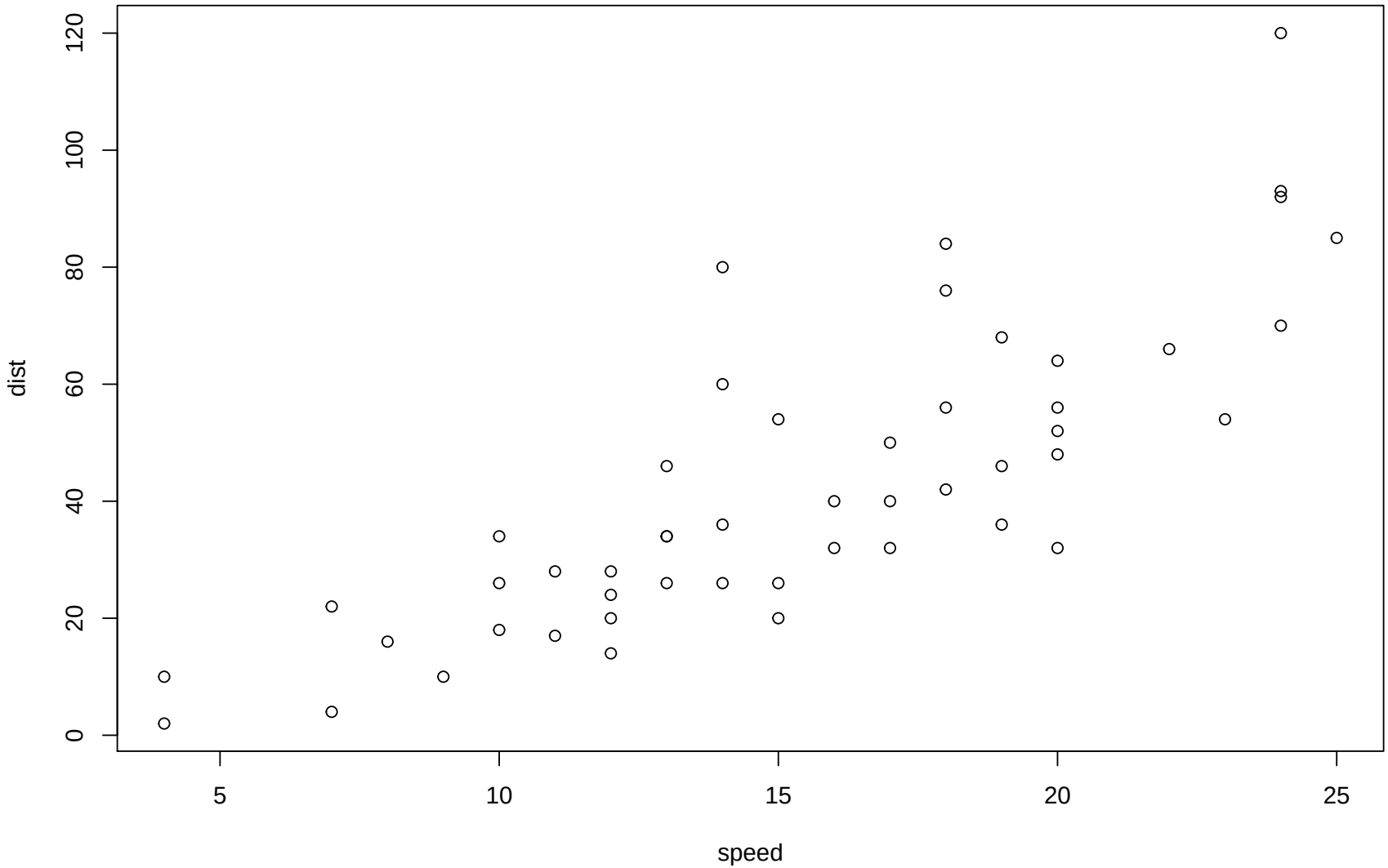
Resources

Slide with R Code and Output

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Slide with Plot



R Markdown

This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see