REVOLUTION ANALYTICS

R at 30,000 ft

Revolution Analytics

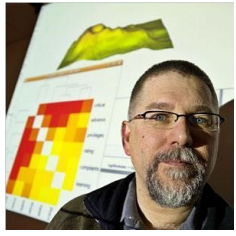
Objectives

- What is R?
- Why should I use it?
- How do I use it?
- Getting help?

Outline

Background and History of R

- R is a successor of S Language
- Originally designed by two University of Auckland Professors for their intro to statistics course.



What is R?

Video at Revolution Analytics blog

R Philosophy

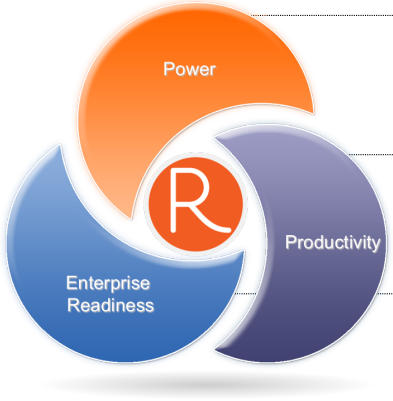
R follows the Unix philosophy

> *Write programs that do one thing and do it well. Write programs to work together.*

- R is extensible with more than 5,000 packages available at CRAN (`http://crantastic.org/packages`)

Revolution Analytics

Revolution R Enterprise scales R to the Enterprise



Power
- Distributed high performance analytics

Productivity
- Build & deploy analytics applications easily

Enterprise Readiness
- Enterprise landscape
- Full-service customer support, consulting and training

CRAN

- CRAN: The Comprehensive R Archive Network
- CRAN is the main repository for R core and its community built libraries.
  - See (http://cran.r-project.org)
- It is the global community of R users and developers that is R's primary strength

Strengths of R

- Expressive
- Open source
- Extendable
- Advanced data structures and graphics
- Large user community
- It is designed by statisticians

Weaknesses of R

- It is designed by statisticians
- Inefficient at element-by-element computations
- May make large demands on system resources, namely memory

Example

```
head(mtcars)
fit0 <- lm(mpg ~ ., data = mtcars)  # fit with all variables
summary(fit0)
library(MASS)
fit1 <- stepAIC(fit0, direction = "both")
summary(fit1)
plot(fit1)
anova(fit1, fit0)
```

Outline

Development environments for R

- Data scientists typically use R inside an integrated development environment (IDE)
- We leverage RStudio in this course

RStudio

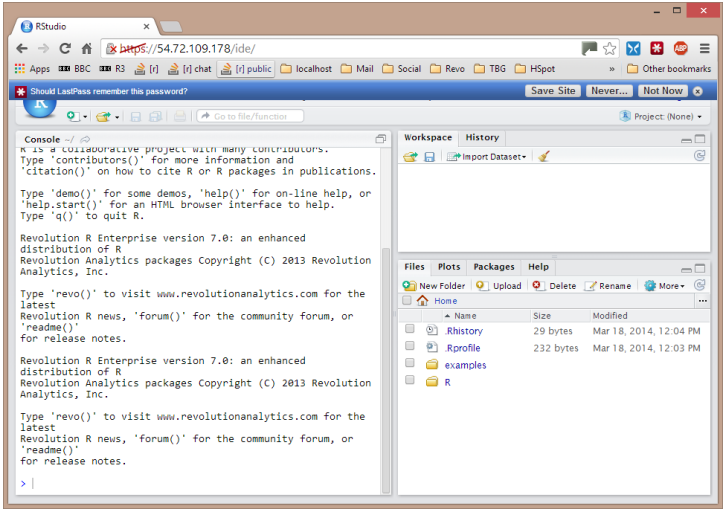


Figure:

Typical IDE

- R terminal
- Script window
- Project files
- Workspace objects

Command line prompts

Symbol	Meaning
>	ready for a new command
+	awaiting the completion of an existing command

Getting help in R

From within the command line:

```
?lm
help(lm)
```

Online help

- StackOverflow
  - http://stackoverflow.com/
- R help mailing list and archives
  - https://stat.ethz.ch/mailman/listinfo/r-help
- Revolution forum
  - http://forums.revolutionanalytics.com/forums/forum.php

Exercise: Get some help

Your turn:

1. Look up the documentation for the `lm` function, which fits an OLS regression to a data set
2. Get together with a partner and come up with a thorny data manipulation, analysis problem, or algorithm that you routinely encounter. Search around the R online community to see who else has been working on that.

Quick review:

- Advantages and disadvantages of R?
- How do you get help?

REVOLUTION ANALYTICS

Thank you

Revolution Analytics is the leading commercial provider of software and support for the popular open source R statistics language.

www.revolutionanalytics.com
1.855.GET.REVO
Twitter: @RevolutionR