Statistical Computing: The Course

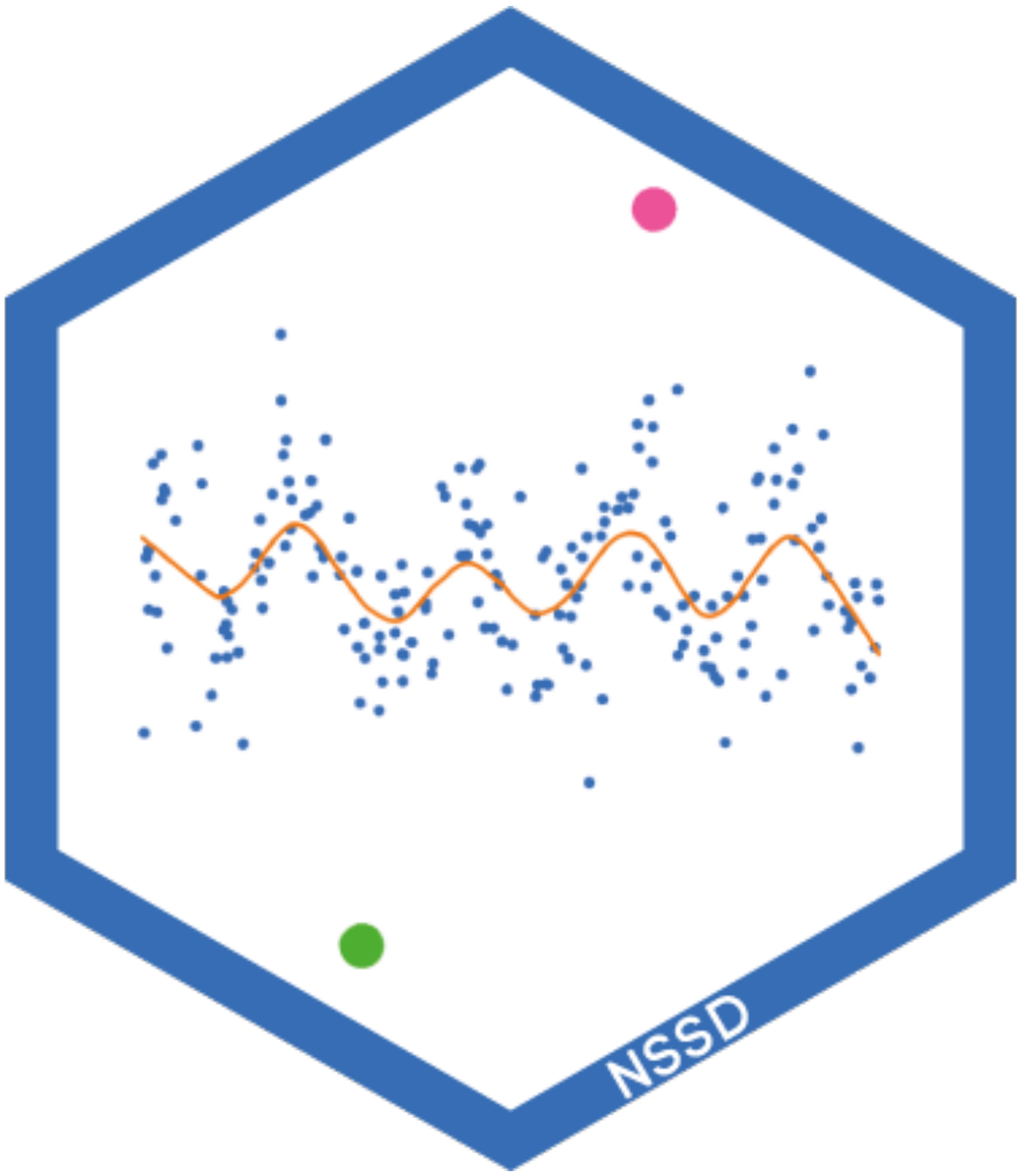
Biostatistics 140.776
Roger D. Peng

About Me

- Outdoor air pollution and health
- Air pollution epidemiology
- Ambient air quality standards
- Time series, spatial statistics, hierarchical modeling

About Me

- Indoor air pollution and health
- Panel studies in vulnerable groups (COPD, asthma)
- Environmental interventions and clinical trials
- Longitudinal data, causal inference, mediation



Not So Standard Deviations

(with Hilary Parker of Stitch Fix)

https://soundcloud.com/nssd-podcast

Subscribe in iTunes: https://goo.gl/ZhWYbd

The Effort Report

(with Elizabeth Matsui of JHU)

http://effortreport.libsyn.com

Subscribe in iTunes: http://goo.gl/hz3AHL

simplystats

simplystatistics.org

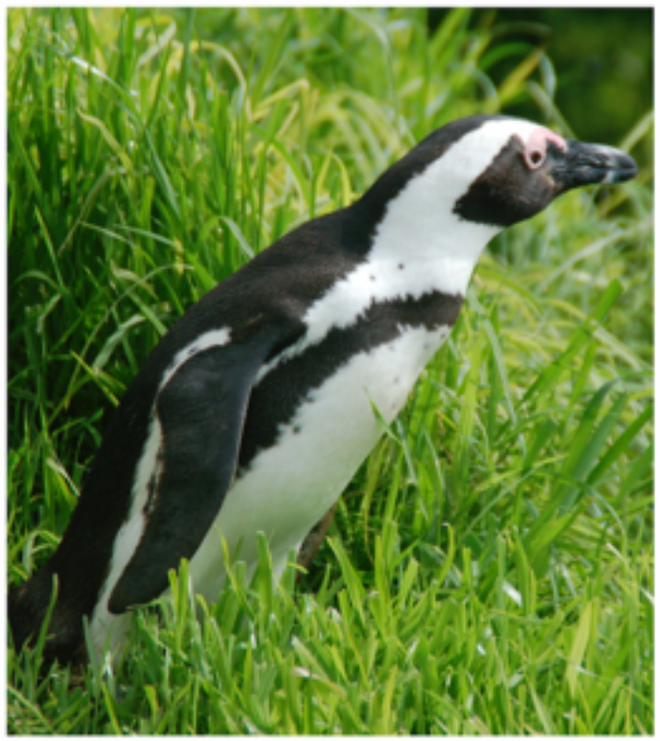
@simplystats

Course Logistics

- **Web Site**: rdpeng.github.io/Biostat776
- **Meets**: T/Th 1:30—2:50pm in W2008
- **Office Hour**: Wednesday 12—1pm (E3535), or email me
- **NOTE**: Lecture on September 15 is **cancelled**; will be recorded instead (I'll send a reminder)

Textbooks

R Programming for Data Science



leanpub.com/rprogramming

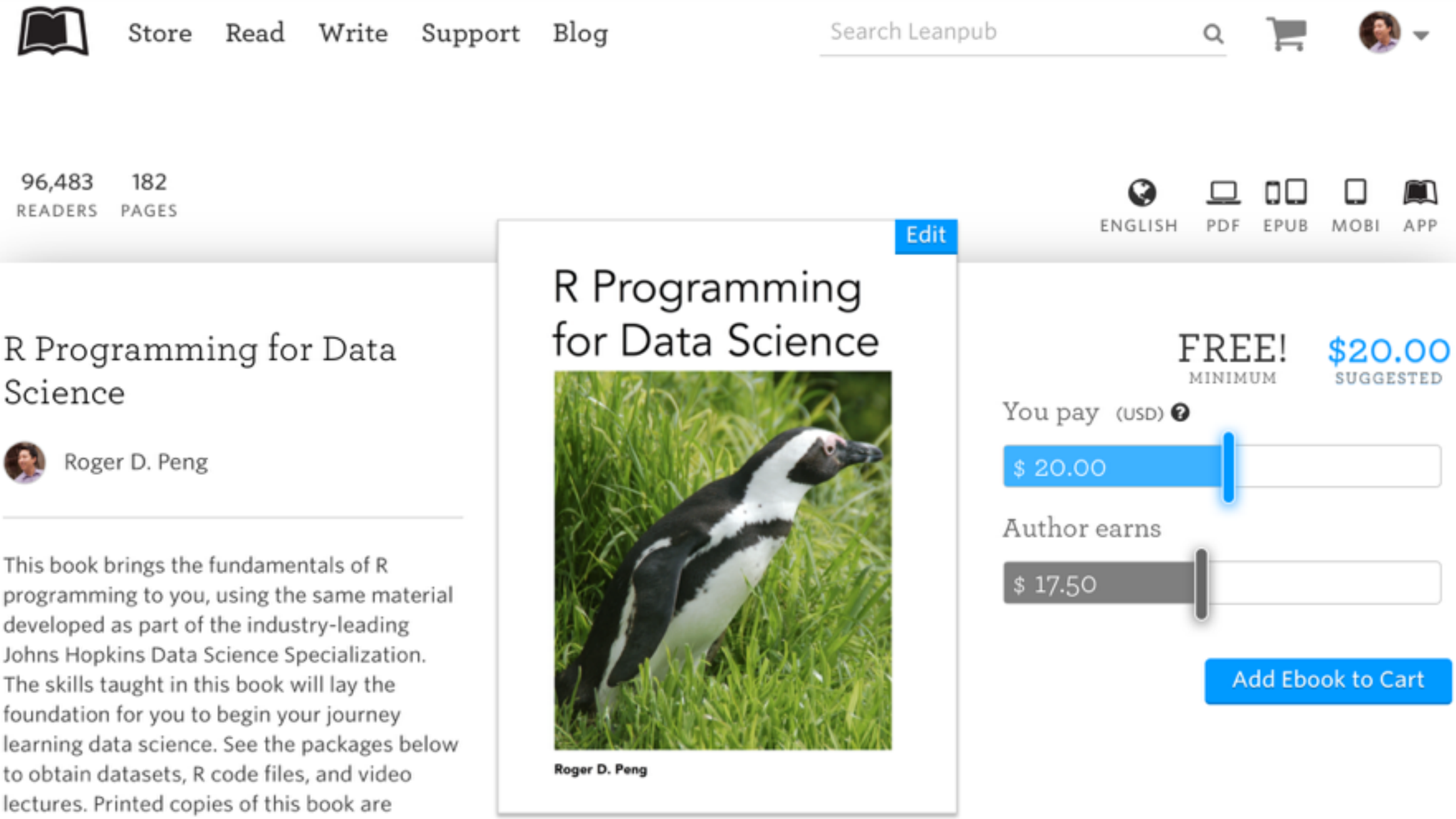
Exploratory Data Analysis with R

leanpub.com/exdata

Report Writing for Data Science in R

leanpub.com/reportwriting

Leanpub


Store
Read
Write
Support
Blog
Search Leanpub
96,483
READERS
182
PAGES
ENGLISH
PDF
EPUB
MOBI
APP
Edit
R Programming for Data Science
Roger D. Peng
UPDATED 28 DAYS AGO
R Programming for Data Science
Roger D. Peng
This book brings the fundamentals of R programming to you, using the same material developed as part of the industry-leading Johns Hopkins Data Science Specialization. The skills taught in this book will lay the foundation for you to begin your journey learning data science. See the packages below to obtain datasets, R code files, and video lectures. Printed copies of this book are available through Lulu.
FREE SAMPLE
FREE!
MINIMUM
$20.00
SUGGESTED
You pay (USD)
$ 20.00
Author earns
$ 17.50
Add Ebook to Cart

Textbooks

R Programming for Research

Colorado State University, ERHS 535

Brooke Anderson and Rachel Severson

geanders.github.io/RProgrammingForResearch/

Mastering Software Development in R

Roger D. Peng, Sean Kross, and Brooke Anderson

2016-08-31

rdpeng.github.io/RProgDA/

Grading

- No exams!
- **Three** homeworks requiring programming in R and some basic data analysis
- Each homework counts equally (1/3)
- **Homeworks submitted via Courseplus dropbox**

Software

- R (of course)
- Various R packages
- You can use whatever you want with respect to Mac, Windows, Linux….
- Obtain R from http://cran.rstudio.com
- I **strongly recommend** RStudio as a nice interface to R (http://www.rstudio.com)

CRAN

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

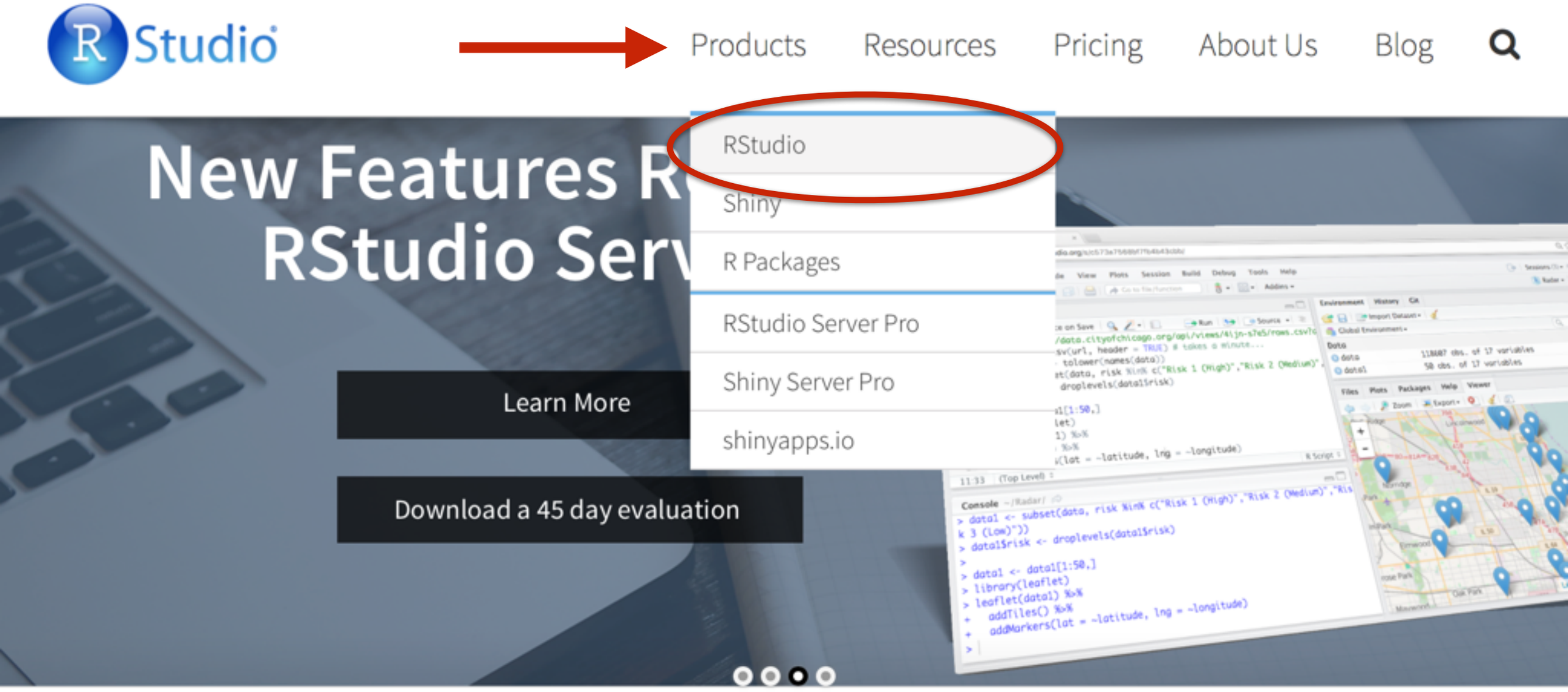
R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

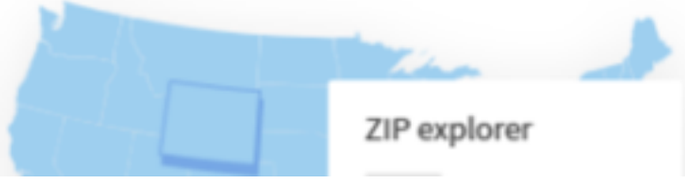
Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (Tuesday 2016-06-21, Bug in Your Hair) R-3.3.1.tar.gz, read what's new in the latest version.
- Sources of R alpha and beta releases (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are available here. Please read about new features and bug fixes before filing corresponding feature requests or bug reports.
- Source code of older versions of R is available here.
- Contributed extension packages

RStudio


RStudio
Products
Resources
Pricing
About Us
Blog
RStudio
Shiny
R Packages
RStudio Server Pro
Shiny Server Pro
shinyapps.io
New Features R
RStudio Serv
Learn More
Download a 45 day evaluation

RSTUDIO
CODE
WORKSPACE


ZIP explorer

rmarkdown
Shiny
tidyr

RStudio

Choose Your Version of RStudio

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. Learn More

	RStudio Desktop (Free License)	RStudio Desktop (Commercial License)	RStudio Server (Free License)	RStudio Server Pro (Commercial License)
Integrated Development Environment for R	✓	✓	✓	✓
Priority support		✓		✓
Access via Web Browser			✓	✓
Enterprise Security and Access Controls				✓
Project Sharing				✓

Major Themes

- Reproducible research
- Data management and manipulation, tidy data
- Data visualization and communication
- Programming with R
- Products and tooling

Major Themes

- knitr, markdown, R markdown
- Tidy data, dplyr, tidyr, lubridate, regular expressions
- Principles of data graphics, ggplot2, mapping
- Functions, functional programming, object oriented programming
- R packages, Shiny apps