

Экономическая
кибернетика:
Введение в
программирование на R

Что такое R и S?

- R- это диалект языка S
- Язык S был разработан Джоном Чемберсом и ко. в Bell Labs
- S появился на свет в 1976 как внутренняя среда статистического анализа, первоначально представлял из себя набор библиотек, написанных на Fortran. Ранние версии языка не включали в себя функции, предназначенные для статистического моделирования.
- В 1988 году система библиотек была переписана на C (эта версия получила номер 3). В это же время появилась книга «Статистические модели в языке S» (Чемберс, Хести).
- Четвертая версия языка S вышла в 1998г., эта та версия, которой мы пользуемся до сих пор.
- Главная идея языка S заключалась в том, что люди должны сначала относиться к нему как к интерактивной среде и уже потом, с опытом переходить на полноценное программирование, погружаясь в языковые и системные аспекты языка. Т.е. должен быть легкий старт. Система должна была помогать пользователю стать программистом.
- R был разработан двумя программистами – Россом Ихака и Робертом Джентлменом в 1991 в Новой Зеландии
- В 1993 R был представлен широкой общественности
- В 1995 Мартин Махлер убеждает Роберта и Росса использовать GNU GPL
- 1997 г. Сформирована группа разработчиков ядра R, исходный код ядра контролируется этой группой. Сторонние разработчики могут предлагать изменения.
- 2000 г. Выпущена версия 1.0.0
- 2013 г. Выпущена версия 3.0.2
- Самая последняя стабильная версия 3.5.1 (июль 2018)

Особенности R

- R – язык программирования для научных вычислений и анализа данных с упором на визуализацию и воспроизводимость
- R – свободное кросс-платформенное программное обеспечение с открытым исходным кодом;
- R является интерпретируемым языком с интерфейсом командной строки;
- R – мультипарадигменный язык, сочетающий в себе:
 1. функциональное
 2. процедурное
 3. объектно-ориентированное
 4. рефлексивное программирование

Особенности R

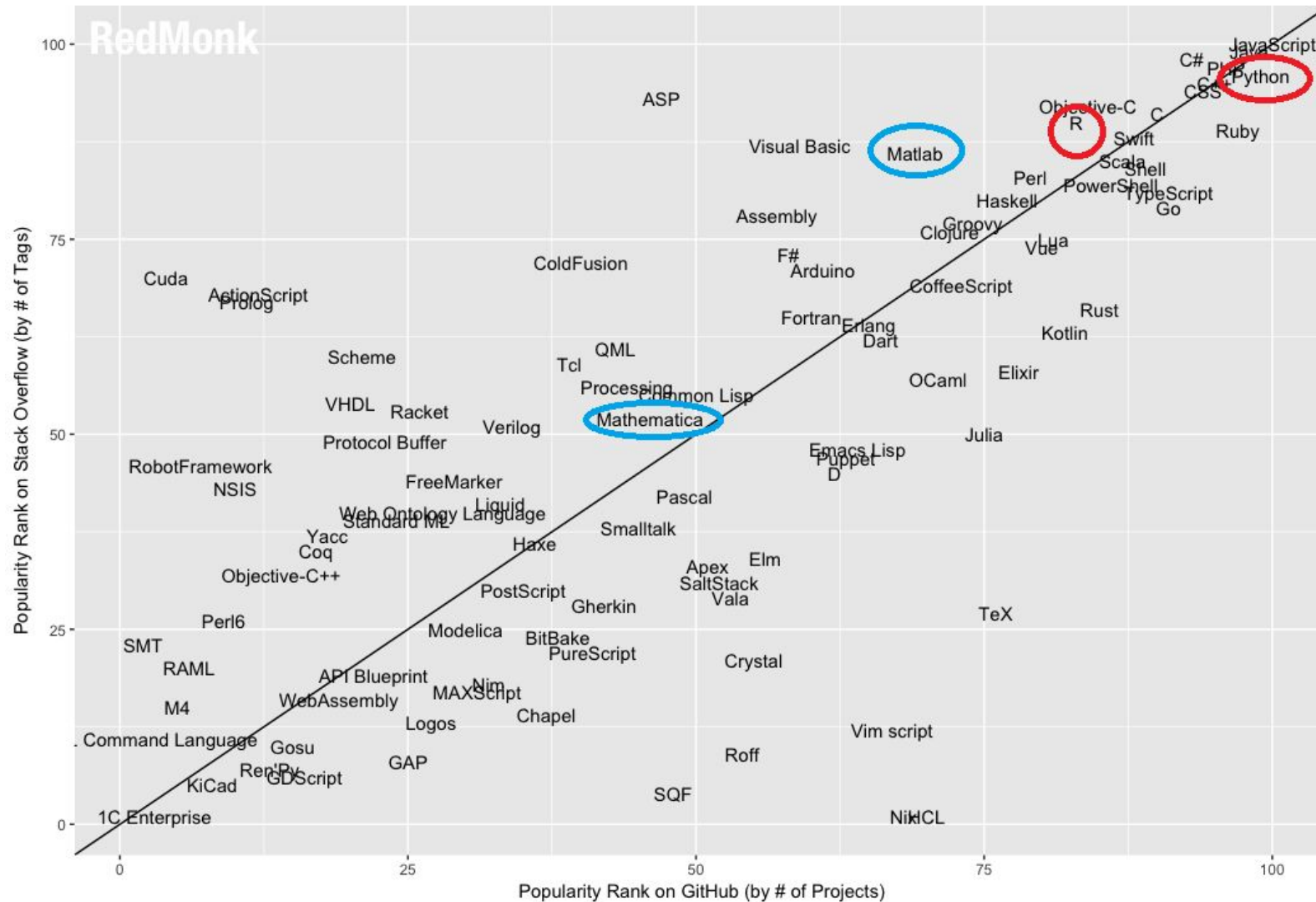
- Синтаксис похож на S;
- Семантика, на первый взгляд тоже похожа, но на деле сильно отличается от S;
- Запускается практически на любой стандартной платформе;
- Частые релизы и постоянное обновление.
- Функционал разбит на множество пакетов, поэтому можно настроить его под себя и не скачивать ненужные пакеты
- Очень сильный графический функционал.
- При этом R подходит не только для интерактивной работы с графикой, в него заложен мощный язык программирования, позволяющий разрабатывать собственные инструменты.
- Очень активное и живое сообщество пользователей доступных через mailing lists или на Stack Overflow.
- Свободность (freedom to run, study, redistribute and improve)

Недостатки R

- Достаточно сложен как язык программирования.
- Документация некоторых методов труднодоступна без глубокого знания математических методов и статистических процедур.
- Легко написать плохой, медленный, неподдерживаемый, нечитаемый код.
- Основан на технологиях 40-летней давности.
- Слабая поддержка динамической и 3d графики.
- Функционал опирается на пользовательский спрос, нет обязательной поддержки, если нужный вам метод не реализован, нет организации, которая брала бы на себя обязанность вам помочь с реализацией этого метода.
- Пакеты дополнений устаревают.
- Не является идеальным для всех возможных задач, как и любой другой язык.

Почему R?

RedMonk Q118 Programming Language Rankings



Альтернативы языки:

- Python,
- Matlab,
- Wolfram Mathematica,
- Octave (совмест. с Matlab)

Альтернативы статистические программы:

- SAS,
- SPSS,
- Statistica,
- Stata,
- Eviews ...

Другие рейтинги

(<https://www.tiobe.com/tiobe-index/>)

Sep 2018	Sep 2017	Change	Programming Language	Ratings	Change
1	1		Java	17.436%	+4.75%
2	2		C	15.447%	+8.06%
3	5	▲	Python	7.653%	+4.67%
4	3	▼	C++	7.394%	+1.83%
5	8	▲	Visual Basic .NET	5.308%	+3.33%
6	4	▼	C#	3.295%	-1.48%
7	6	▼	PHP	2.775%	+0.57%
8	7	▼	JavaScript	2.131%	+0.11%
9	-	▲▲	SQL	2.062%	+2.06%
10	18	▲▲	Objective-C	1.509%	+0.00%
11	12	▲	Delphi/Object Pascal	1.292%	-0.49%
12	10	▼	Ruby	1.291%	-0.64%
13	16	▲	MATLAB	1.276%	-0.35%
14	15	▲	Assembly language	1.232%	-0.41%
15	13	▼	Swift	1.223%	-0.54%
16	17	▲	Go	1.081%	-0.49%
17	9	▼▼	Perl	1.073%	-0.88%
18	11	▼▼	R	1.016%	-0.80%
19	19		PL/SQL	0.850%	-0.63%
20	14	▼▼	Visual Basic	0.682%	-1.07%

SPSS, SAS, R

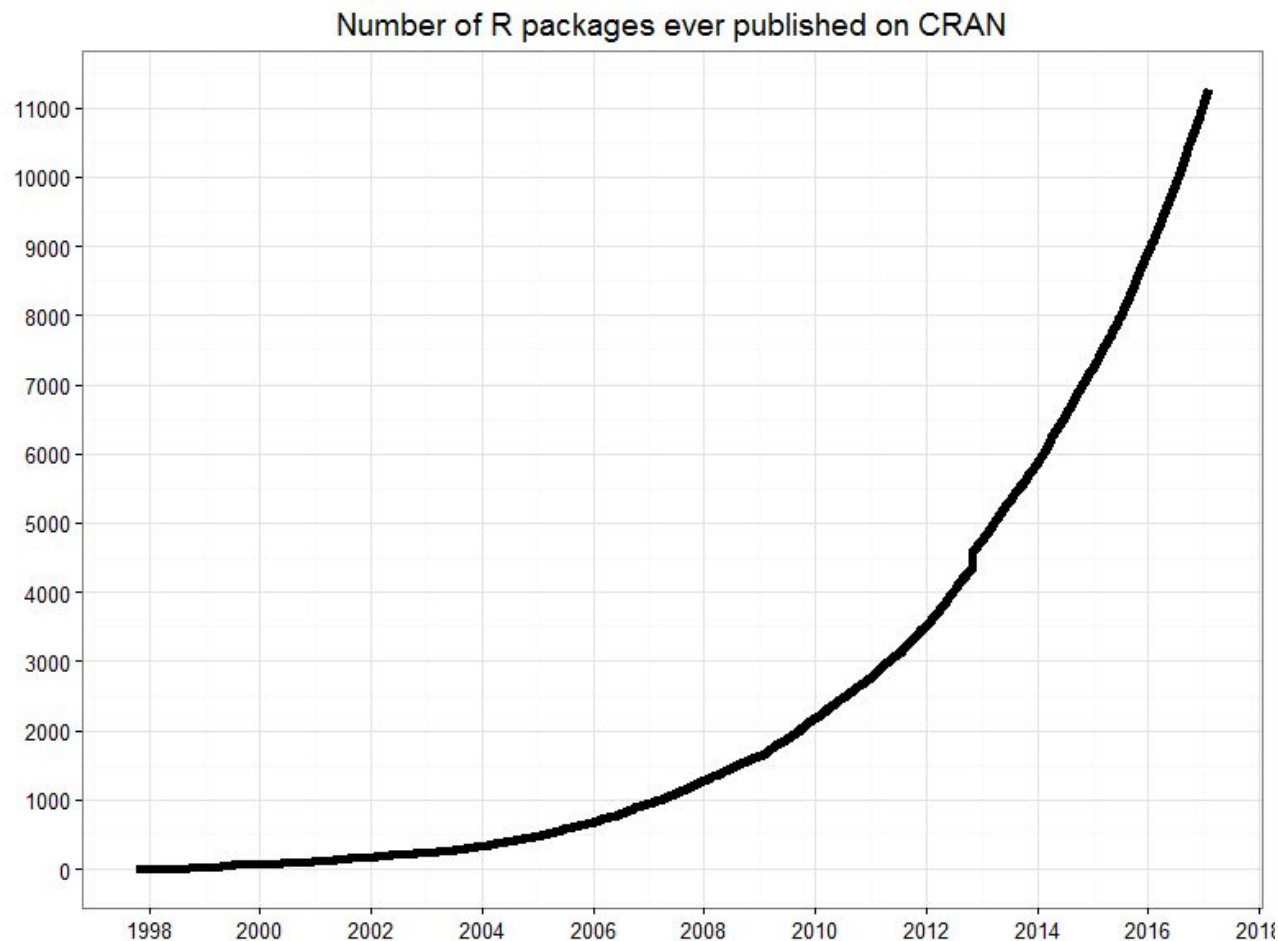
- SPSS -> Revolution analytics
- В январе 2015 Microsoft приобретает Revolution analytics
- SAS – считается более надежным, используется крупными компаниями Boeing, РЖД и т.д.

Python и R:

- Пока библиотеки R богаче,
- Чаще всего библиотеки пишутся академ. сообществом, методы реализованы надежно, но не всегда оптимальны
- Python быстрее,
- Python сокращает отставание.

R:

- Бесплатный,
- Популярный,
- Доступный,
- Проблема с памятью.



Источник:

<http://blog.revolutionanalytics.com/2017/01/cran-10000.html>

Структура среды R

Язык R разбит на две концептуальные части:

1. «базовая» часть, которую мы загружаем из репозитория CRAN;
2. все остальное

Функционал языка разделен на множество пакетов:

- «базовая» часть включает в себя пакет **base**, который содержит все основные функции низшего уровня, которые вам необходимы при работе в R,
- Другие пакеты, содержащиеся в «базовой» части включают: **utils, stats, datasets, graphics, grDevices, grid, methods, tools, parallel, compiler, splines, tcltk, stats4**.
- Также есть перечень “Рекомендованных” пакетов: **boot, class, cluster, codetools, foreign, KernSmooth, lattice, mgcv, nlme, rpart, survival, MASS, spatial, nnet, Matrix**.

The Comprehensive R Archive Network

← → ↻ https://cran.r-project.org/index.html



The Comprehensive R Archive Network

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](#)

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 (2018-07-02, Feather Spray) [R-3.5.1.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](#)

Questions About R

- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

What are R and CRAN?

R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc. Please consult the [R project homepage](#) for further information.

CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R. Please use the CRAN [mirror](#) nearest to you to minimize network load.

Submitting to CRAN

To "submit" a package to CRAN, check that your submission meets the [CRAN Repository Policy](#) and then use the [web form](#).

If this fails, upload to <ftp://CRAN.R-project.org/incoming/> and send an email to CRAN-submissions@R-project.org following the policy. Please do not attach submissions to emails, because this will clutter up the mailboxes of half a dozen people.

Note that we generally do not accept submissions of precompiled binaries due to security reasons. All binary distribution listed above are compiled by selected maintainers, who are in charge for all binaries of their platform, respectively.


For queries about this web site, please contact [the webmaster](#).

Специализированные пакеты (Cran Task Views)

Пакеты сгруппированы по областям знаний / задачам:

Bayesian	Bayesian Inference	Multivariate	Multivariate Statistics
ChemPhys	Chemometrics and Computational Physics	NaturalLanguageProcessing	Natural Language Processing
ClinicalTrials	Clinical Trial Design, Monitoring, and Analysis	NumericalMathematics	Numerical Mathematics
Cluster	Cluster Analysis & Finite Mixture Models	OfficialStatistics	Official Statistics & Survey Methodology
DifferentialEquations	Differential Equations	Optimization	Optimization and Mathematical Programming
Distributions	Probability Distributions	Pharmacokinetics	Analysis of Pharmacokinetic Data
Econometrics	Econometrics	Phylogenetics	Phylogenetics, Especially Comparative Methods
Environmetrics	Analysis of Ecological and Environmental Data	Psychometrics	Psychometric Models and Methods
ExperimentalDesign	Design of Experiments (DoE) & Analysis of Experimental Data	ReproducibleResearch	Reproducible Research
ExtremeValue	Extreme Value Analysis	Robust	Robust Statistical Methods
Finance	Empirical Finance	SocialSciences	Statistics for the Social Sciences
FunctionalData	Functional Data Analysis	Spatial	Analysis of Spatial Data
Genetics	Statistical Genetics	SpatioTemporal	Handling and Analyzing Spatio-Temporal Data
Graphics	Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization	Survival	Survival Analysis
HighPerformanceComputing	High-Performance and Parallel Computing with R	TimeSeries	Time Series Analysis
MachineLearning	Machine Learning & Statistical Learning	WebTechnologies	Web Technologies and Services
MedicalImaging	Medical Image Analysis	gR	gRaphical Models in R
MetaAnalysis	Meta-Analysis		
ModelDeployment	Model Deployment with R		

RStudio



The image shows a screenshot of the RStudio website homepage. At the top, there is a navigation bar with the RStudio logo and links for Products, Resources, Pricing, About Us, and Blogs. The main content area features a large blue banner with the RStudio logo and the text "Open source and enterprise-ready professional software for R". Below this banner are four buttons: "Download RStudio", "Discover Shiny", "shinyapps.io Login", and "Discover RStudio Connect". The lower section of the page is divided into three columns, each representing a different RStudio component: RStudio, Shiny, and R Packages. Each column includes a representative image, a title, a brief description, and a "Learn More" link.

← → ↻ https://www.rstudio.com

R Studio

Products Resources Pricing About Us Blogs

RStudio

Open source and enterprise-ready professional software for R

Download RStudio

Discover Shiny

shinyapps.io Login

Discover RStudio Connect



RStudio

RStudio makes R easier to use. It includes a code editor, debugging & visualization tools.

Download Learn More



Shiny

Shiny helps you make interactive web applications for visualizing data. Bring R data analysis to life.

Learn More



R Packages

Our developers create popular packages to expand the features of R. Includes ggplot2, dplyr, R Markdown & more.

Learn More

Rstudio GUI

The screenshot displays the RStudio interface with the following components:

- Source Editor:** A script editor window titled "Untitled1*" containing a single line of R code: `1`.
- Environment Pane:** Shows the "Global Environment" with two variables: `a` and `b`, both containing the value `"Hello world"`.
- Console:** Displays the R prompt and the output of the executed code:

```
~$R/ > print("Hello world")
[1] "Hello world"
> a<-print("Hello world")
[1] "Hello world"
> a
[1] "Hello world"
> b=a
> b
[1] "Hello world"
> |
```
- Help Viewer:** Shows the documentation for the `fBasics` package, titled "Portfolio Modelling, Optimization and Backtesting". The description states: "The Rmetrics 'fBasics' package is a collection of functions to explore and to investigate basic properties of financial returns and related quantities. The covered fields include techniques of explorative data analysis and the investigation of distributional properties, including parameter estimation and hypothesis testing. Evenmore there are several utility functions for data handling and managemnet." Details include: Package: `fBasics`, Type: `Package`, Date: `2014`, License: `GPL Version 2 or later`, Copyright: `(c) 1999-2014 Rmetrics Association`, Repository: `R-FORGE`, URL: `https://www.rmetrics.org`.

Если нужна помощь

1. `> help("lm")`
2. `> help.search("logarithm")`
3. R-FAQ на CRAN
4. Vignettes
5. Google
6. Stackoverflow.com (сначала ищем ответ, потом спрашиваем)
7. R Mailing List Archive
8. Package maintainer's e-mail.