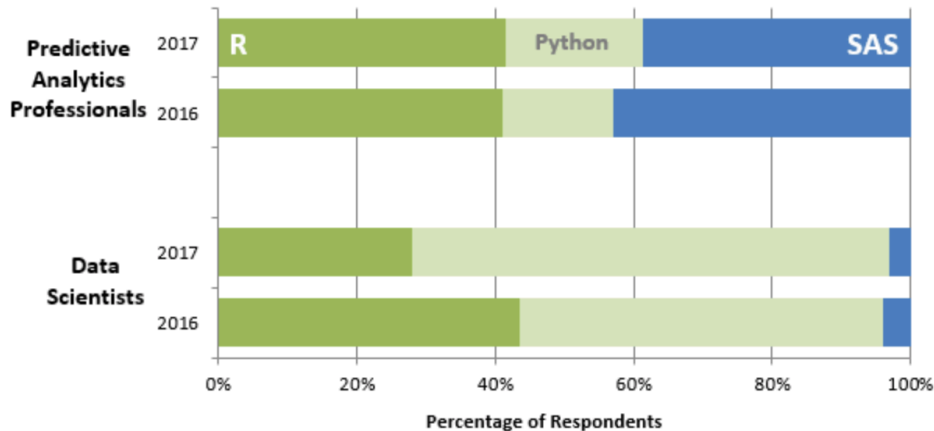


**Getting started with R**

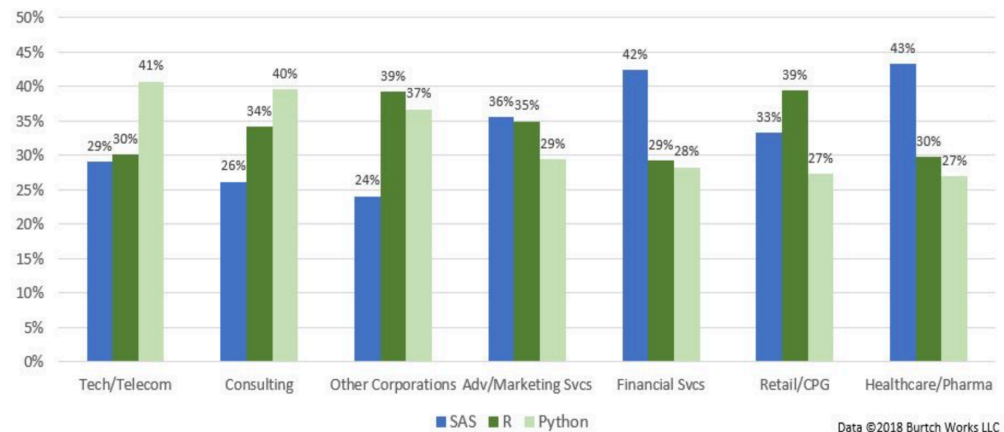
# How popular is R?

R is a good statistical computing language

**Tool Preference for Data Scientists vs. Other Predictive Analytics Professionals**



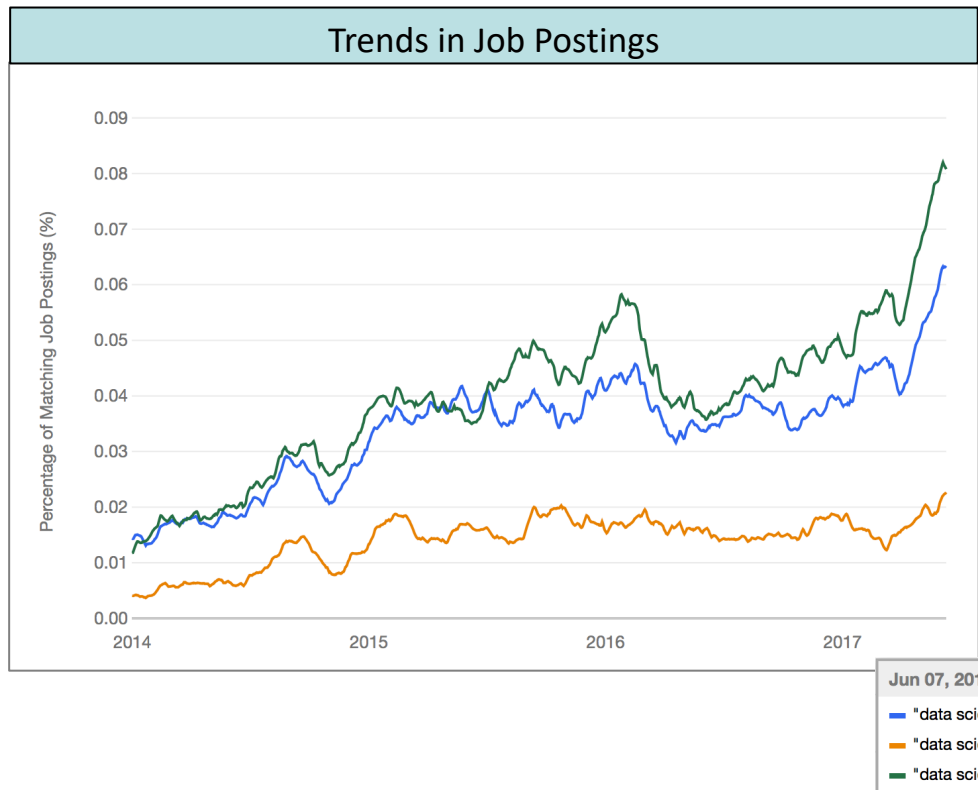
**SAS, R, or Python Preference by Industry**



Data ©2018 Burtch Works LLC

# How popular is R?

## R and data science in the job market



Salary estimated from 7,098 employees, users, and past and present job advertisements on Indeed in the past 36 months. Last updated: February 2, 2019

### Location

United States

### Average salary

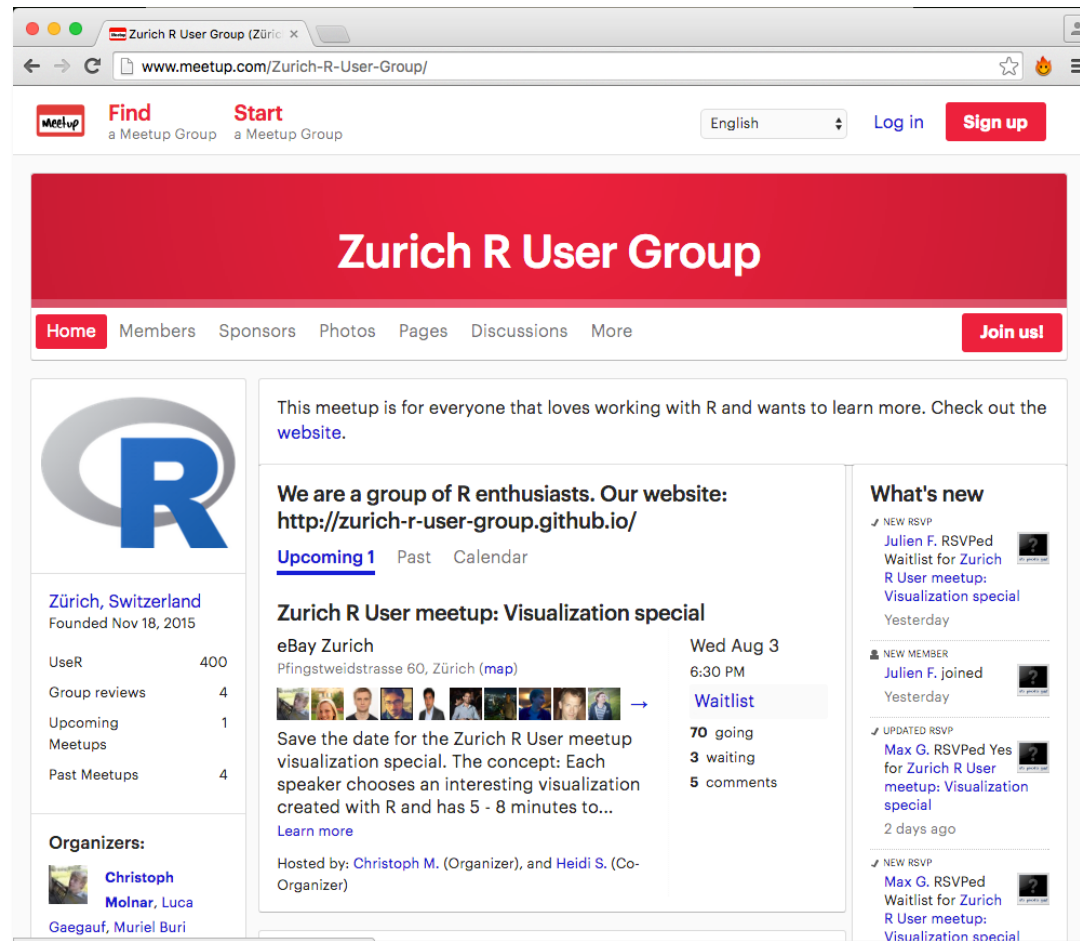
**\$127,981** per  
year

### Salary Distribution



# How popular is R?

## There is also a local R user community in Zurich



The screenshot shows the Meetup page for the Zurich R User Group. The page has a red header with the group name "Zurich R User Group" and navigation links like "Home", "Members", "Sponsors", "Photos", "Pages", "Discussions", and "More". A "Join us!" button is also present. The main content area features a large "R" logo, a description of the group, and details about an upcoming "Visualization special" meetup. The sidebar on the left provides statistics and organizer information.

**Zurich R User Group**

Home Members Sponsors Photos Pages Discussions More **Join us!**

This meetup is for everyone that loves working with R and wants to learn more. Check out the [website](#).

We are a group of R enthusiasts. Our website: <http://zurich-r-user-group.github.io/>

**Upcoming 1** Past Calendar

**Zurich R User meetup: Visualization special**

eBay Zurich  
Pfingstweidstrasse 60, Zürich ([map](#))

Wed Aug 3  
6:30 PM

[Waitlist](#)

70 going  
3 waiting  
5 comments

Save the date for the Zurich R User meetup visualization special. The concept: Each speaker chooses an interesting visualization created with R and has 5 - 8 minutes to... [Learn more](#)

Hosted by: [Christoph M.](#) (Organizer), and [Heidi S.](#) (Co-Organizer)

**Zürich, Switzerland**  
Founded Nov 18, 2015

Category	Count
UserR	400
Group reviews	4
Upcoming Meetups	1
Past Meetups	4

**Organizers:**

- [Christoph Molnar](#), Luca Gaegauf, Muriel Buri

**What's new**

- NEW RSVP: [Julien F.](#) RSVPed Waitlist for Zurich R User meetup: Visualization special Yesterday
- NEW MEMBER: [Julien F.](#) joined Yesterday
- UPDATED RSVP: [Max G.](#) RSVPed Yes for Zurich R User meetup: Visualization special 2 days ago
- NEW RSVP: [Max G.](#) RSVPed Waitlist for Zurich R User meetup: Visualization special

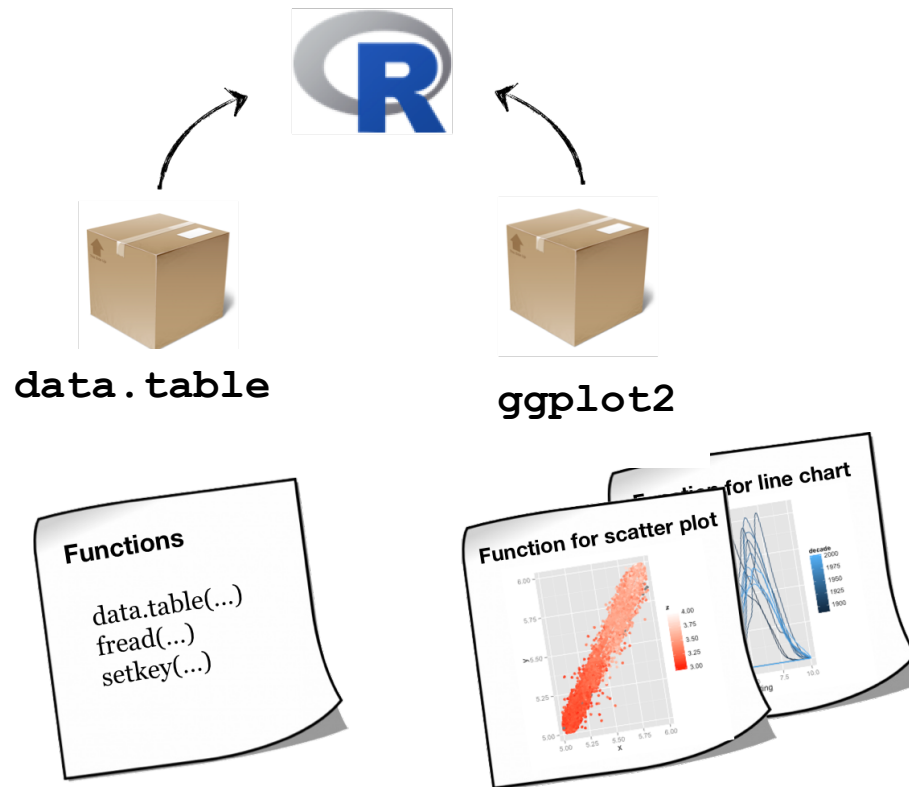
# Why is R so popular?

## There are many reasons...

- R is open and free.
- RStudio is a great IDE for R and also open and free.
- R runs on many operating systems (Windows, Mac OS X, Linux, Unix).
- R is easily extensible via user-developed packages.
- R is scalable.
- Analyses done using R are reproducible.
- Using R makes collaboration easier.
- Finding answers to questions is simple as the R community is very helpful.

# Why is R so popular?

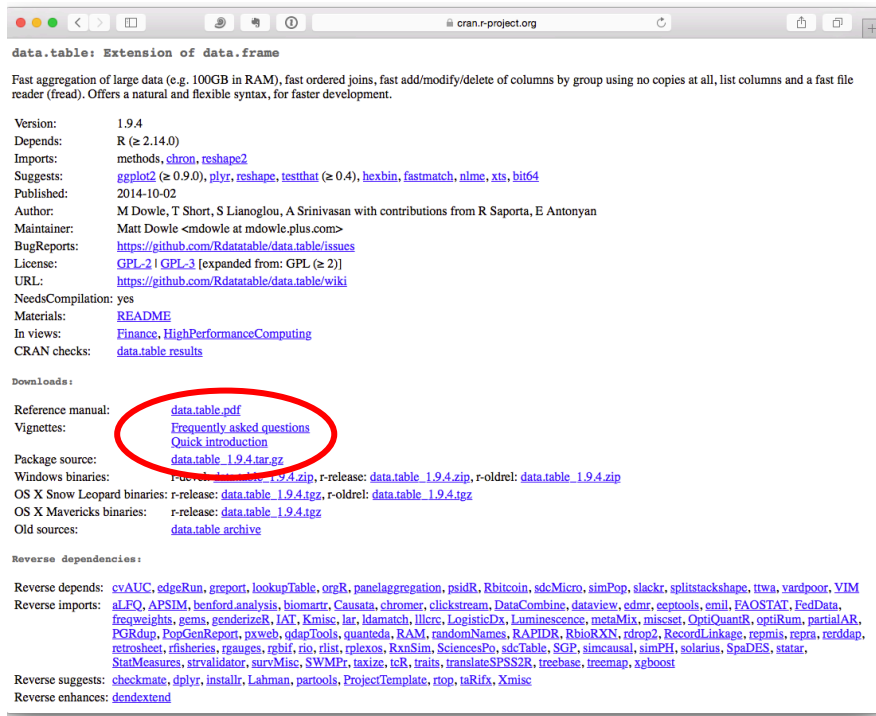
## R packages extend the functionality of R



- **Packages** are collections of R functions, data, and compiled code in a well-defined format.
- They are downloaded mostly from **CRAN** (Comprehensive R Archival Network).
- <https://cran.r-project.org>
- In total, more than 16'000 packages are available from CRAN (10/2020).

# Why is R so popular?

## In most cases, good documentation of R packages



```

data.table: Extension of data.frame

Fast aggregation of large data (e.g. 100GB in RAM), fast ordered joins, fast add/modify/delete of columns by group using no copies at all, list columns and a fast file reader (fread). Offers a natural and flexible syntax, for faster development.

Version: 1.9.4
Depends: R (≥ 2.14.0)
Imports: methods, chron, reshape2
Suggests: ggplot2 (≥ 0.9.0), plyr, reshape, testthat (≥ 0.4), hexbin, fastmatch, nlm, xts, bit64
Published: 2014-10-02
Author: M Dowle, T Short, S Lianoglou, A Srinivasan with contributions from R Saporta, E Antonyan
Maintainer: Matt Dowle <mdowle at mdowle.plus.com>
BugReports: https://github.com/Rdatatable/data.table/issues
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL: https://github.com/Rdatatable/data.table/wiki
NeedsCompilation: yes
Materials: README
In views: Finance, HighPerformanceComputing
CRAN checks: data.table results

Downloads:

Reference manual: data.table.pdf
Vignettes: Frequently asked questions
Quick introduction
Package source: data.table_1.9.4.tar.gz
Windows binaries: r-release: data.table_1.9.4.zip, r-oldrel: data.table_1.9.4.zip
OS X Snow Leopard binaries: r-release: data.table_1.9.4.tgz, r-oldrel: data.table_1.9.4.tgz
OS X Mavericks binaries: r-release: data.table_1.9.4.tgz
Old sources: data.table archive

Reverse dependencies:

Reverse depends: cvAUC, edgeRun, greport, lookupTable, orgR, panelaggregation, psidR, Rbitcoin, sdcMicro, simPop, slackr, splitstackshape, ttw, vardpoor, YIM
alFO, APSIM, benford.analysis, biomatr, Causata, chromer, clickstream, DataCombine, dataview, edmr, reptools, emli, FAOSTAT, FedData,
freeweight, gams, gendertizeR, IAT, Kmisc, lar, idamatch, llicr, LogisticDx, Luminescence, metaMix, miscat, OptiQuantR, optiRun, partialAR,
PGRdup, PopGenReport, pxweb, qdapTools, quanteda, RAM, randomNames, RAPIDR, RbioRXN, rdrop2, RecordLinkage, repmis, repra, renddap,
retrosheet, rfisheries, rgauges, rgbif, rio, rlist, rplexos, RxnSim, SciencesPo, sdcTable, SGP, simcausal, simPH, solarius, SpaDES, statar,
StatMeasures, survvalidator, survMisc, SYMPr, taxize, tcr, traits, translateSPSS2R, treebase, treeemap, xgboost
Reverse suggests: checkmate, dplyr, installr, Lahman, partools, ProjectTemplate, rtop, taRifx, Xmisc
Reverse enhances: dendextend
  
```

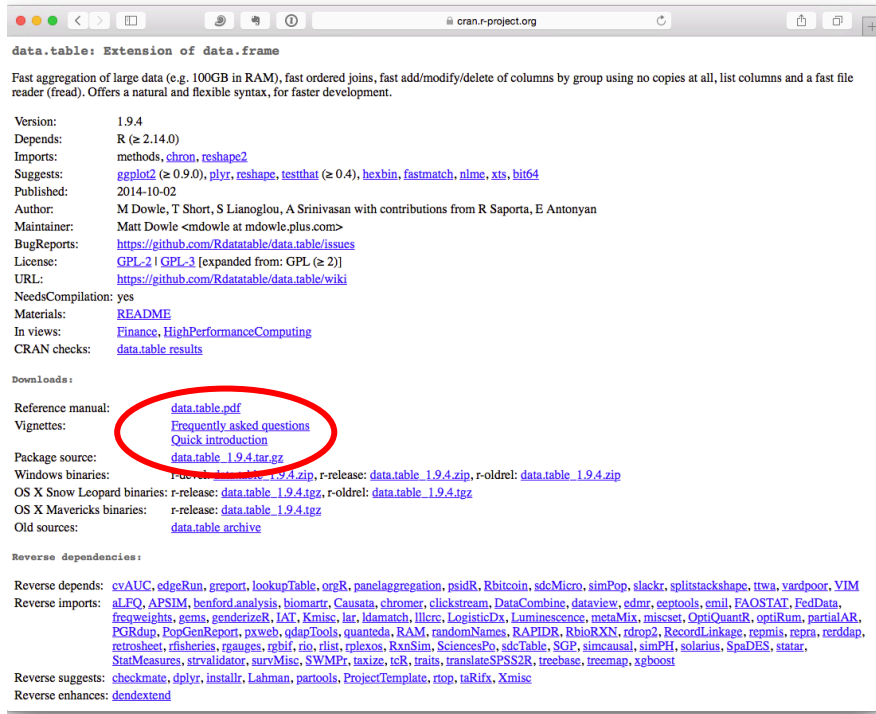
Check, for example, if a walkthrough or package vignette exists, either on CRAN or by using the R command

```
browseVignettes("package.name")
```

after installing the R package.

# Why is R so popular?

## In most cases, good documentation of R packages



```

data.table: Extension of data.frame

Fast aggregation of large data (e.g. 100GB in RAM), fast ordered joins, fast add/modify/delete of columns by group using no copies at all, list columns and a fast file reader (fread). Offers a natural and flexible syntax, for faster development.

Version: 1.9.4
Depends: R (≥ 2.14.0)
Imports: methods, chron, reshape2
Suggests: ggplot2 (≥ 0.9.0), plyr, reshape, testthat (≥ 0.4), hexbin, fastmatch, nlm, xts, bit64
Published: 2014-10-02
Author: M Dowle, T Short, S Lianoglou, A Srinivasan with contributions from R Saporta, E Antonyan
Maintainer: Matt Dowle <mdowle at mdowle.plus.com>
BugReports: https://github.com/Rdatatable/data.table/issues
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL: https://github.com/Rdatatable/data.table/wiki
NeedsCompilation: yes
Materials: README
In views: Finance, HighPerformanceComputing
CRAN checks: data.table results

Downloads:

Reference manual: data.table.pdf
Vignettes: Frequently asked questions
Quick introduction
Package source: data.table_1.9.4.tar.gz
Windows binaries: r-release: data.table_1.9.4.zip, r-oldrel: data.table_1.9.4.zip
OS X Snow Leopard binaries: r-release: data.table_1.9.4.tgz, r-oldrel: data.table_1.9.4.tgz
OS X Mavericks binaries: r-release: data.table_1.9.4.tgz
Old sources: data.table archive

Reverse dependencies:

Reverse depends: cvAUC, edgeRun, greport, lookupTable, orgR, panelaggregation, psidR, Rbitcoin, sdcMicro, simPop, slackr, splitstackshape, ttw, vardpoor, YIM
alFO, APSIM, benford.analysis, biomatr, Causata, chromer, clickstream, DataCombine, dataview, edmr, eeptools, emli, FAOSTAT, FedData,
freeweight, gams, gendertizeR, IAT, Kmisc, lar, idamatch, ilicr, LogisticDx, Luminescence, metaMix, miscat, OptiQuantR, optiRun, partialAR,
PGRdup, PopGenReport, pxweb, qdapTools, quanteda, RAM, randomNames, RAPIDR, RbioRXN, rdrop2, RecordLinkage, repmis, repa, renddap,
retrosheet, rfisheries, rgauges, rgbif, rio, rlist, rplexos, RxnSim, SciencesPo, sdcTable, SGP, simcausal, simPH, solarius, SpaDES, statar,
StatMeasures, survvalidator, survMisc, SYMPr, taxize, tcr, traits, translateSPSS2R, treebase, treeap, xgboost
Reverse suggests: checkmate, dplyr, installr, Lahman, partools, ProjectTemplate, rtop, taRifx, Xmisc
Reverse enhances: dendextend
  
```

Check, for example, if a walkthrough or package vignette exists, either on CRAN or by using the R command

```
browseVignettes("package.name")
```

after installing the R package.



# Basics of documentation in R

## 1) Package documentation

**Package ‘data.table’**

April 7, 2019

**Version** 1.12.2

**Title** Extension of ‘data.frame’

**Depends** R (>= 3.1.0)

**Imports** methods

**Suggests** bit64, curl, R.utils, knitr, xts, nanotime, zoo

**Description** Fast aggregation of large data (e.g. 100GB in RAM), fast ordered joins, fast add/modify/delete of columns by group using no copies at all, list columns, friendly and fast character-separated-value read/write. Offers a natural and flexible syntax, for faster development.

**License** MPL-2.0 | file LICENSE

**URL** <http://r-datatable.com>

- Gives an overview of the package and explains all functions
- Format: PDF file
- Typically displayed on the CRAN page  
<https://cran.r-project.org/web/packages/>

## 2) Documentation of functions

R Documentation

**cbind (base)**

Combine R Objects by Rows or Columns

**Description**

Take a sequence of vector, matrix or data-frame arguments and combine by columns or rows, respectively. These are generic functions with methods for other R classes.

**Usage**

```
cbind(..., deparse.level = 1)
rbind(..., deparse.level = 1)
## S3 method for class 'data.frame'
rbind(..., deparse.level = 1, make.row.names = TRUE,
       stringsAsFactors = default.stringsAsFactors(), factor.exclude = NA)
```

**Arguments**

... (generalized) vectors or matrices. These can be given as named arguments. Other R objects may be coerced as appropriate, or S4 methods may be used: see sections ‘Details’ and ‘Value’. (For the “data.frame” method of cbind these can be further arguments to [data.frame](#) such as `stringsAsFactors`.)

- All functions should have a help page in the R documentation
- They can be assessed via  
`help(functionname)` or  
`?functionname`

# Basics of documentation in R

## 1) Package documentation

**Package ‘data.table’**

April 7, 2019

**Version** 1.12.2

**Title** Extension of ‘data.frame’

**Depends** R (>= 3.1.0)

**Imports** methods

**Suggests** bit64, curl, R.utils, knitr, xts, nanotime, zoo

**Description** Fast aggregation of large data (e.g. 100GB in RAM), fast ordered joins, fast add/modify/delete of columns by group using no copies at all, list columns, friendly and fast character-separated-value read/write. Offers a natural and flexible syntax, for faster development.

**License** MPL-2.0 | file LICENSE

**URL** <http://r-datatable.com>

- Gives an overview of the package and explains all functions
- Format: PDF-file
- Typically displayed on the CRAN page  
<https://cran.r-project.org/web/packages/>

## 2) Documentation of functions

R Documentation

**cbind (base)**

Combine R Objects by Rows or Columns

**Description**

Take a sequence of vector, matrix or data-frame arguments and combine by columns or rows, respectively. These are generic functions with methods for other R classes.

**Usage**

```
cbind(..., deparse.level = 1)
rbind(..., deparse.level = 1)
## S3 method for class 'data.frame'
rbind(..., deparse.level = 1, make.row.names = TRUE,
       stringsAsFactors = default.stringsAsFactors(), factor.exclude = NA)
```

**Arguments**

... (generalized) vectors or matrices. These can be given as named arguments. Other R objects may be coerced as appropriate, or S4 methods may be used: see sections ‘Details’ and ‘Value’. (For the “data.frame” method of cbind these can be further arguments to [data.frame](#) such as `stringsAsFactors`.)

- All functions should have a help page in the R documentation
- They can be assessed via  
`help(functionname)` or  
`?functionname`

**Getting started with R**