



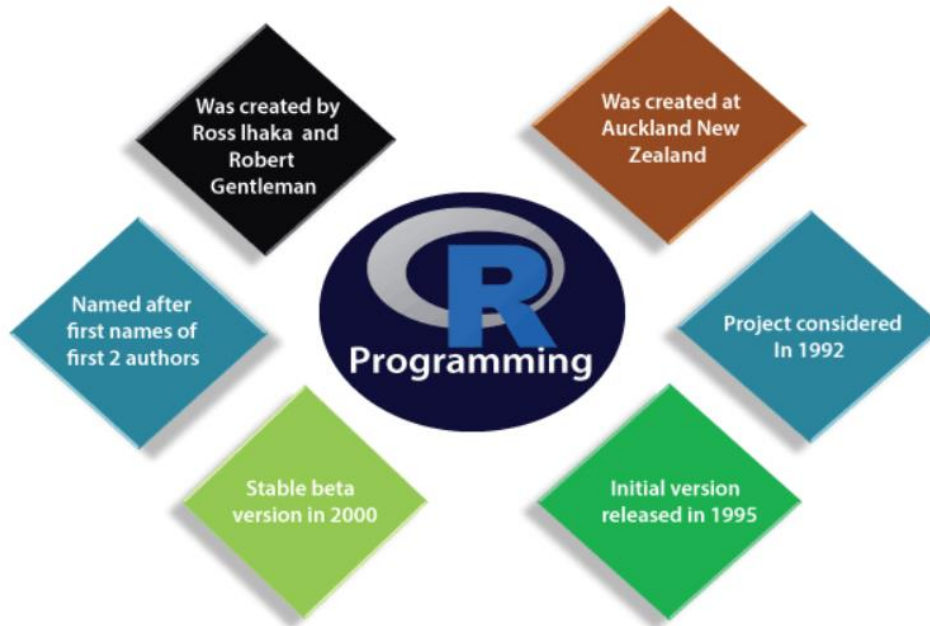
R

Programming

R Programming

- **"R is an interpreted computer programming language which was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand."**
- **R is software environment used to analyze statistical information, graphical representation, reporting, and data modeling.**

History of R Programming

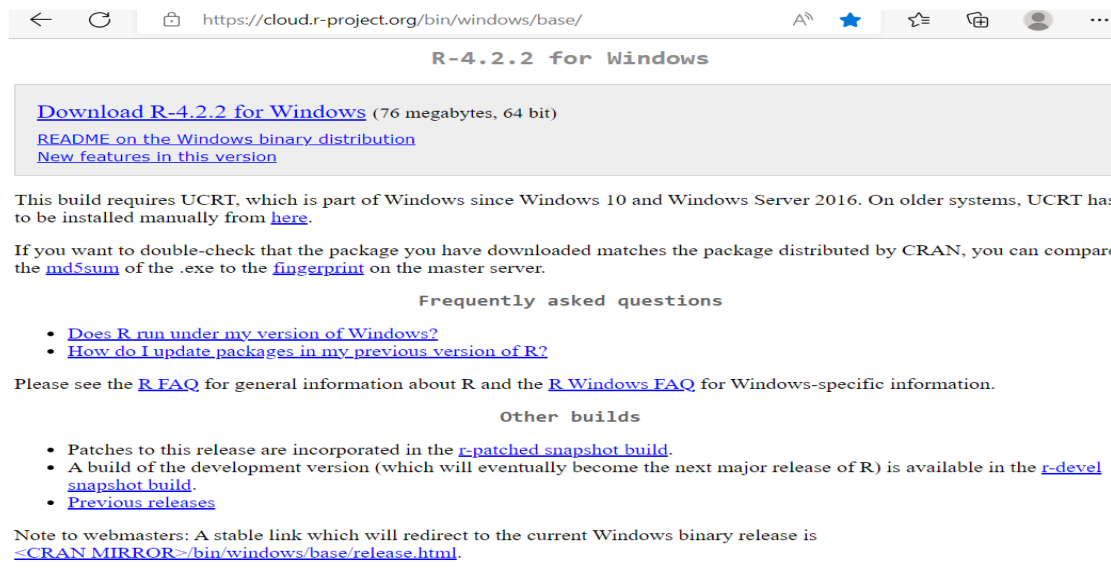


Features of R

- It is a simple and effective programming language which has been well developed.
- It is data analysis software.
- **It is an open-source, powerful, and highly extensible software.**
- **It provides highly extensible graphical techniques.**

Installation of R

- Go to the site <https://cloud.r-project.org/bin/windows/base/>



The screenshot shows a web browser window with the address bar displaying <https://cloud.r-project.org/bin/windows/base/>. The page title is "R-4.2.2 for Windows". The main content area has a light gray background and contains the following text:

[Download R-4.2.2 for Windows](#) (76 megabytes, 64 bit)
[README on the Windows binary distribution](#)
[New features in this version](#)

This build requires UCRT, which is part of Windows since Windows 10 and Windows Server 2016. On older systems, UCRT has to be installed manually from [here](#).

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

Note to webmasters: A stable link which will redirect to the current Windows binary release is [-CRAN_MIRROR>/bin/windows/base/release.html](https://cloud.r-project.org/bin/windows/base/release.html).

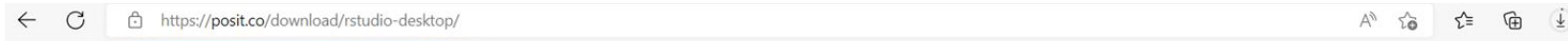
Contd..

- Select the path where we want to download the R and proceed to Next.
- Select all components which we want to install, and then we will proceed to Next.
- In the next step, we have to select either customized startup or **accept the default**, and then we proceed to Next.
- When we proceed to next, our installation of R in our system will get started:
- In the last, we will click on finish to successfully install R in our system.

RStudio IDE

- RStudio is an integrated development environment which allows us to interact with R
- Go to the website <https://posit.co/download/rstudio-desktop/>

Rstudio IDE



PRODUCTS ▾

SOLUTIONS ▾

LEARN & SUPPORT ▾

EXPLORE MORE ▾



DOWNLOAD RSTUDIO

Step 2: Install RStudio Desktop

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 190.49MB | [SHA-256: B38BF925](#) | Version: 2022.07.2+576 |

Released: 2022-09-21

Exploring Gui

The screenshot displays the RStudio graphical user interface (GUI) with the following components:

- Source Editor:** Contains a script named `firstR.R` with the following R code:

```
1 string<-"Hello world!"
2 print(string)
```
- Console:** Shows the R version and system information:

```
R 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```
- Environment:** Displays "Environment is empty".
- Files:** Shows a file explorer view of the `PG-DBDA` project directory. The file list includes:

Name	Size	Modified
assignment 5.docx	11.9 KB	Oct 31, 2022, 5:13 PM
Data Warehousing.pptx	103 KB	Oct 1, 2022, 12:54 PM
Data Wrangling With Python.mp4	72.9 MB	Oct 24, 2022, 10:28 AM
database-mcq-day1.docx	18.3 KB	Sep 26, 2022, 4:29 PM
database-mcq-day1.pdf	106.1 KB	Sep 26, 2022, 4:33 PM
Database-mcq-day5.docx	11.6 KB	Oct 1, 2022, 2:45 PM
Database-mcq-day6.docx	16.9 KB	Oct 1, 2022, 3:23 PM
Dataframe-questions.docx	12 KB	Oct 27, 2022, 4:24 PM
Day-6-Python		
dd.txt	5 B	Oct 15, 2022, 3:43 PM
f1.html	480 B	Nov 1, 2022, 10:17 AM
firstsite.html	204 B	Nov 1, 2022, 9:47 AM
firstsite.txt	113 B	Nov 1, 2022, 9:40 AM
ggplot.pptx	38.7 KB	Oct 31, 2022, 2:54 PM
java-assignment.xlsx	12.3 KB	Nov 17, 2022, 12:56 PM
Java-mcq-5.docx	21.8 KB	Nov 8, 2022, 10:54 PM
Java-mcq-day9-10.docx	21.9 KB	Nov 13, 2022, 9:44 PM
Java-mcq-day13.docx	23 KB	Nov 17, 2022, 11:12 AM

R studio

RStudio Windows/Tabs	Location	Description
Console Window	Lower-left	The location where commands are entered and output is printed.
Source Tabs	Upper-left	Built-in text editor
Environment Tab	Upper-left	An interactive list of loaded R objects.
History Tab	Upper-left	List of keystrokes entered into the console.
Files Tab	Lower-right	File explorer to navigate C drive folders.
Plots Tab	Lower-right	Output location for plots.
Packages Tab	Lower-right	List of installed packages.
Help Tab	Lower-right	Output location for help commands and help search Window.
Viewer Tab	Lower-right	Advanced tab for local web content.

R Packages

- R packages are the collection of R functions, sample data, and compile codes.
- In the R environment, these packages are stored under a directory called "library."
- During installation, R installs a set of packages.
- We can add packages later when they are needed for some specific purpose.

R Packages

- R provides `library()` function, which allows us to get the list of all the installed packages.
- R provides `search()` function to get all packages currently loaded in the R environment.
- To install a new package- **`install.packages("Package Name")`**

Contd..



Contd..

- **Tidyr**-The word tidyr comes from the word tidy, which means clear. So the tidyr package is used to make the data 'tidy'.
- **Ggplot2**- This package is famous for its elegant and quality graphs which sets it apart from other visualization packages.
- **Ggraph** - R provides an extension of ggplot known as ggraph.
- **Dplyr** - R allows us to perform data wrangling and data analysis.
- **Tidyquant** -The tidyquant is a financial package which is used for carrying out quantitative financial analysis.

Contd..

- Dygraphs - The dygraphs package provides an interface to the main JavaScript library which we can use for charting.
- Leaflet - For creating interactive visualization, R provides the leaflet package.
- Ggmap - It is a mapping package which consists of various tools for geolocating and routing.
- Glue - R provides the glue package to perform the operations of data wrangling.

Contd ..

- Shiny -R allows us to develop interactive and aesthetically pleasing web apps by providing a shiny package.
- Plotly -The plotly package provides online interactive and quality graphs.
- **Tidytex**- The tidytext package provides various functions of text mining for word processing and carrying out analysis through ggplot, dplyr, and other miscellaneous tools.
- Stringr- string operations

Contd..

- Reshape2- This package facilitates flexible reorganization and aggregation of data
- Dichromat -The R dichromat package is used to remove Red-Green or Blue-Green contrasts from the colors.
- Digest - The digest package is used for the creation of cryptographic hash objects of R functions.
- MASS - The MASS package provides a large number of statistical functions.

Contd..

- caret - R allows us to perform classification and regression tasks by providing the caret package.
- E1071 - The e1071 library provides useful functions which are essential for data analysis like Naive Bayes, Fourier Transforms, SVMs, Clustering, and other miscellaneous functions.
- sentiment - The sentiment package provides functions for carrying out sentiment analysis

Syntax of R

- `string <- "Hello World!"`
- `print(string)`
- Comments in R- Single line and multiline

Variables in R Programming

Variables are used to store the information to be manipulated and referenced in the R program.

R is a dynamically typed, means it check the type of data type when the statement is run.

A valid variable name contains letter, numbers, dot and underlines characters.

A variable name should start with a letter or the dot not followed by a number.

_var_name	Invalid	Variable name can't start with an underscore(_).
var_name, var.name	Valid	Variable can start with a dot, but dot should not be followed by a number. In this case, the variable will be invalid.
var_name%	Invalid	In R, we can't use any special character in the variable name except dot and underscore.
2var_name	Invalid	Variable name cant starts with a numeric digit.
.2var_name	Invalid	A variable name cannot start with a dot which is followed by a digit.
var_name2	Valid	The variable contains letter, number and underscore and starts with a letter.

Assignment of variable

In R programming, there are three operators which we can use to assign the values to the variable.

We can use leftward, rightward, and equal_to operator for this purpose.

R allows you to assign the same value to multiple variables in one line

Output functions

There are two functions which are used to print the value of the variable i.e., `print()` and `cat()`.

The `cat()` function combines multiples values into a continuous print output.

join, two or more elements, by using the `paste()` function

Data Types in R Programming

Data type	Example	Description
Logical	True, False	It is a special data type for data with only two possible values which can be construed as true/false.
Numeric	12,32,112,5432	Decimal value is called numeric in R, and it is the default computational data type.
Integer	3L, 66L, 2346L	Here, L tells R to store the value as an integer,
Complex	Z=1+2i, t=7+3i	A complex value in R is defined as the pure imaginary value i.
Character	'a', '"good"', "TRUE", '35.4'	In R programming, a character is used to represent string values. We convert objects into character values with the help of as.character() function.
Raw		A raw data type is used to hold raw bytes.

Keywords

In programming, a keyword is a word which is reserved by a program because it has a special meaning.

A keyword can be a command or a parameter.

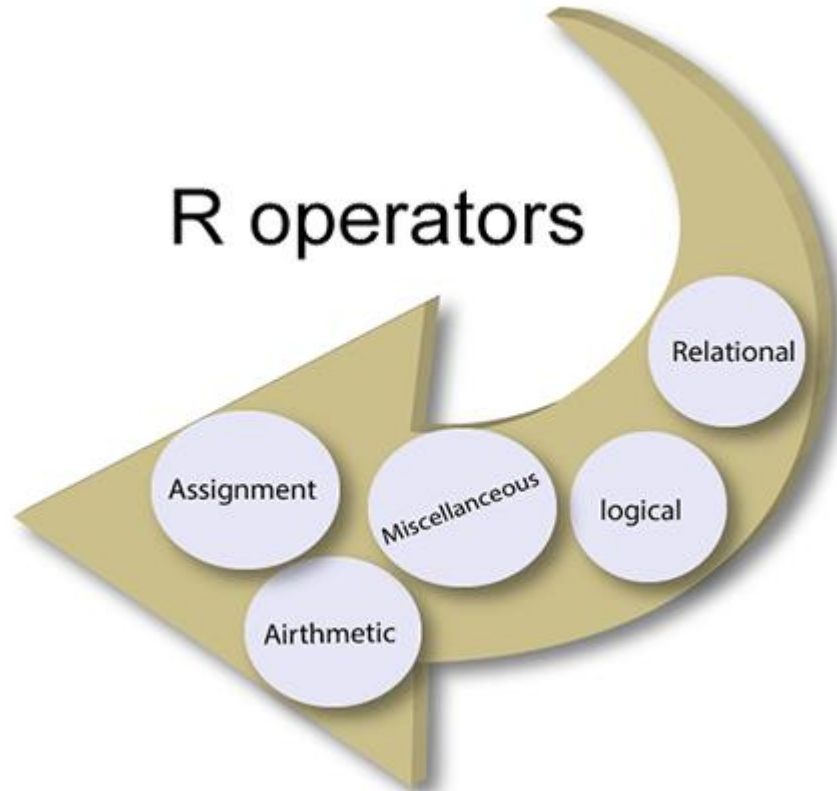
A keyword can't be used as a variable name.

Keywords are also called as "reserved names."

Keywords



Operators in R



Arithmetic Operators

Arithmetic operators are the symbols which are used to represent arithmetic math operations.

`+, -, *, /`

`%%` This operator is used to find the remainder

`/%` This operator is used to find the quotient

Relational Operators

A relational operator is a symbol which defines some kind of relation between two entities.

These include numerical equalities and inequalities.

Logical Operators

The logical operators allow a program to make a decision on the basis of multiple conditions.

In the program, each operand is considered as a condition which can be evaluated to a false or true value.

The value of the conditions is used to determine the overall value of the op1 **operator** op2.

& This operator is known as the Logical AND operator.

| This operator is called the Logical OR operator.

! This operator is known as Logical NOT operator.

&& This operator compares variables and gives TRUE as a result, only if both are TRUE.

|| This operator compares variables and gives the result TRUE, if one of them is true.

Assignment Operators

An assignment operator is used to assign a new value to a variable.

<- or = or <<- These operators are known as left assignment operators.

-> or ->> These operators are known as right assignment operators.

Miscellaneous Operators

Miscellaneous operators are used for a special and specific purpose. These operators are not used for general mathematical or logical computation.

: The colon operator is used to create the series of numbers in sequence for a vector.

%in% This is used when we want to identify if an element belongs to a vector.

%% It is used to multiply a matrix with its transpose.

R if Statement

The if statement consists of the Boolean expressions followed by one or more statements.

The if statement is the simplest decision-making statement which helps us to take a decision on the basis of the condition.

Switch in R

A switch statement is a selection control mechanism that allows the value of an expression to change the control flow of program execution via map and search.

The switch statement is used in place of long if statements which compare a variable with several integral values.

It is a multi-way branch statement which provides an easy way to dispatch execution for different parts of code.

This code is based on the value of the expression.

R next Statement

The next statement is used to skip any remaining statements in the loop and continue executing.

In simple words, a next statement is a statement which skips the current iteration of a loop without terminating it.

R while loop

A while loop is a type of control flow statements which is used to iterate a block of code several numbers of times.

The while loop terminates when the value of the Boolean expression will be false.

```
while (test_expression) {  
    statement  
}
```

R For Loop

A for loop is the most popular control flow statement.

A for loop is used to iterate a vector.

It is similar to the while loop.

There is only one difference between for and while, i.e., in while loop, the condition is checked before the execution of the body, but in for loop condition is checked after the execution of the body.

```
for (value in vector) {  statements  }
```