



Introducción a R para OBIS

Julián Pizarro y Andrea Polanco

Instituto de Investigaciones Marinas y Costeras – INVEMAR

Centro Regional de Entrenamiento temas marinos para Latinoamérica - OTGA-IODE-IOC-UNESCO)

CURSO ADMINISTRACIÓN DE DATOS BIOGEOGRÁFICOS (CONTRIBUYENDO AL USO DE OBIS)

Junio 12 al 16 de 2017

Santa Marta, Colombia



Except where otherwise noted, content in this work is licensed under a [Creative Commons Attribution-ShareAlike 3.0 \(CC BY SA\)](https://creativecommons.org/licenses/by-sa/3.0/) (IGO)

With the support of the
Government of Flanders,
Belgium



Flanders
State of the Art



Que es R?

- R es un lenguaje de programación para estadísticos
 - Es un modo de comunicarnos con las computadoras para que hagan cálculos
- Es un lenguaje vectorial
- ¿Por qué usar R?
 - Es gratis
 - Multitud de utilidades en bioinformática
 - Aceptada por la comunidad científica



Entorno de programación

Instalación de R: <http://cran.freestatistics.org/>

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 (Friday 2017-04-21, You Stupid Darkness) [R-3.4.0.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.



Entorno de programación

RSTUDIO: <https://www.rstudio.com/products/rstudio/download/>

The screenshot displays the RStudio IDE interface. The top-left pane shows a script file named 'test.R' with the following code:

```
1 a <- seq(1, 10)
2
3
```

The bottom-left pane shows the console output:

```
> a <- seq(1, 10)
> ?seq
> |
```

The top-right pane shows the Environment tab with 'Global Environment' selected. The 'Values' section displays a table:

| Variable | Value |
|----------|---------------------------------|
| a | int [1:10] 1 2 3 4 5 6 7 8 9 10 |

The bottom-right pane shows the R Documentation for 'seq {base}'. The title is 'Sequence Generation'. The description states: 'Generate regular sequences. seq is a standard generic with a default method. seq.int is a primitive which can be much faster but has a few restrictions. seq_along and seq_len are very fast primitives for two common cases.' The usage section shows:

```
seq(...)
```

Default S3 method:
seq(from = 1, to = 1, by = ((to - from)/(length.out - 1)),



Paquetes R

Librerías de código reutilizable

La instalación de librerías se realiza de la siguiente manera

CRAN

```
install.packages("ggplot2")
```

GitHub

```
install.packages("devtools")
```

```
devtools::install_github("ropensci/rgbif")
```



Paquetes R

Las librerías más utilizadas para OBIS son:

| | |
|---|---|
| | |
| RObis | <code>install.packages("devtools")</code> <code>devtools::install_github("iobis/robis")</code> |
| Shapefile | <code>install.packages("rgdal")</code> <code>install.packages("mapproj")</code> <code>install.packages("rgeos")</code> <code>install.packages("rmapshaper")</code> |
| Manipular archivos Excel | <code>install.packages("xlsx")</code> |
| Transformar datos tabulares | <code>install.packages("dplyr")</code> |
| Obtener datos de http://www.marineregions.org | <code>install.packages("mregions")</code> |
| | |



Tipos de datos

Vectores: es la estructura de datos mas simple que utiliza R, una colección ordenada de números.

```
> a <- c(1, 2)
> a
[1] 1 2
> b <- seq(1, 10)
> b
[1] 1 2 3 4 5 6 7 8 9 10
> length(b)
[1] 10
```



Tipos de datos

Matrices: estructura de datos de dos dimensiones. Todos los datos son de la misma clase

```
> matrix(1:6, nrow=3, ncol=2)
  [,1] [,2]
[1,]  1  4
[2,]  2  5
[3,]  3  6
```




Tipos de datos

Hojas de datos (Data frames): matrices cuyas columnas pueden tener diferentes modos y atributos

```
> d = data.frame(a=c(1, 2, 3), b=c("x", "y", "z"))
  a b
1 1 x
2 2 y
3 3 z
> d$a
[1] 1 2 3
> d[1]
  a
1 1
2 2
3 3
> d[1,]
  a b
1 1 x
> d[,1]
[1] 1 2 3
> d[["a"]]
[1] 1 2 3
> d[, "a"]
[1] 1 2 3
```



Tipos de datos

Listas (Lists): en una colección ordenada de objetos, conocidos como componentes.

```
Lst <- list(nombre="Pedro", esposa="Maria", no.hijos=3, edad.hijos=c(4,7,9))
```

Los componentes de una lista pueden tener nombre, en cuyo caso pueden ser referidos también por dicho nombre, mediante una expresión de la forma nombre de ***lista\$nombre*** de componente



Tipos de datos

Listas (Lists): en una colección ordenada de objetos, conocidos como componentes.

```
> a <- data.frame(a = c(1, 2, 3), b = c("x", "y", "z"))
> l <- list(a = a, b = 1)
> l$a
a b
1 1 x
2 2 y
3 3 z
> l$b
[1] 1
```



Leer datos

Textos delimitados:

```
data <- read.table("data.txt", header=TRUE, sep="\t", dec=".", stringsAsFactors=FALSE)
data <- read.csv("data.csv")
```

Archivos de excel

```
require(xlsx)

data <- read.xlsx("data.xlsx", 1)
data <- read.xlsx("data.xlsx", sheetName="somesheet")
```



Leer datos

Shapefile:

```
require(maptools)
require(rgdal)

download.file("http://iobis.org/geoserver/OBIS/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=OBIS:summaries&outputFormat=SHAPE-ZIP", destfile="summaries.zip")
unzip("summaries.zip")

shape <- readOGR("summaries.shp", layer="summaries")
```



Referencias

Manual de R

<https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf>

<http://www.iobis.org/manual/intror/>