

PRÁCTICA 1

ENTORNO DE PROGRAMACIÓN R Y SU APLICACIÓN EN EL ANÁLISIS CUANTITATIVO EN EPIDEMIOLOGÍA

1. PRESENTACIÓN

El **R** es un entorno de programación, análisis estadístico y software gráfico derivado del lenguaje de programación **S** y **SPlus** (Becker, Chambers y Wilks, 1988; Chambers, 1998; Chambers y Hastie, 1992; Venables y Ripley, 2000). La primera versión del R fue desarrollada en el Departamento de Estadística de la Universidad de Auckland (Nueva Zelanda) por Ross Ihaka y Robert Gentleman (Ihaka y Gentleman, 1996).

Desde su creación, el R se alimenta y crece con los trabajos de investigadores y docentes provenientes de universidades de todo el mundo y prácticamente todas las ramas del conocimiento. Las aportaciones continuas y desinteresadas de funciones y paquetes de propósito general o específico perfilan a R como un entorno dinámico formado por una comunidad activa y adherida a la filosofía del software libre.

R, es un software libre que está comprendido dentro del proyecto GNU (*General Public Licence* (Licencia Pública General, GNU). Se trata de una licencia creada por la *Free Software Foundation* (Fundación para el software libre) organización fundada por Richard Matthew Stallman en el año 1985. El principal propósito de la licencia GNU es declarar la libertad del uso, modificación y distribución del software y protegerlo de intentos de privatización que puedan de algún modo restringir su uso (el contenido de la licencia puede consultarse en el sitio <https://www.gnu.org/>).

Dentro de esta licencia se distribuyen una serie de programas, muchos de los cuales son versiones libres del software privativo generalista más utilizado. De entre ellos tal vez los más extendidos sean la suite ofimática OpenOffice, el navegador Mozilla, los artículos de wikipedia, el sistema operativo GNU/Linux, o el editor de textos Emacs.

El R integra una multitud de **paquetes** cuya continua incorporación al entorno R incrementan su capacidad y versatilidad. Más adelante será definido que son los paquetes, podríamos equipararlos con los módulos que ofrece el software privativo y comercial para el análisis de datos. El R dispone de funciones básicas relacionadas con los análisis descriptivos de datos, y de los modelos más complejos y actuales concernientes con los últimos avances en el campo de la estadística, la psicometría, la econometría o el análisis de datos en áreas como la agronomía y sus ramas, psicología, economía, sociología, estadística, biología, enfermería, farmacia, medicina o informática.

Aparte de las capacidades de análisis estadístico, R es un potentísimo generador de gráficos. Permite componer un gráfico simple, definir figuras extremadamente complejas e incluso crear animaciones (Maindonald y Braun, 2007; Murrel, 2005; Sarkar, 2008).

2. OBJETIVOS DE APRENDIZAJE

En la presente práctica el estudiante aprenderá lo siguiente:

- 2.1 Instalar el entorno de programación R y el entorno de desarrollo integrado RStudio.
- 2.2 Identificar las ventanas del RStudio y su descripción.
- 2.3 Conocer los principales conceptos en programación en R
- 2.4 Utilizar el R como calculadora.
- 2.5 Comprender los tipos de objetos en R y su construcción
- 2.6 Instalar paquetes con funciones estadísticas específicas (fBasics, agricolae, readxl, entre otros).
- 2.7 Realizar análisis exploratorio de datos usando funciones del R y paquetes

3. DURACIÓN DE LA PRÁCTICA

La práctica 1 tendrá una duración de 4 horas presenciales y 2 virtuales de trabajo autónomo.

4. DESARROLLO DE LA PRÁCTICA

4.1 Para la instalación del R y RStudio siga los siguientes pasos:

a) Instalar R:

Navegar a <http://cran.r-project.org/bin/windows/base/> y

Instale la versión 4.3.1 (ingrese a **Previous releases**) y luego de clic en R-4.3.1-win.exe

R-4.3.3 for Windows

[Download R-4.3.3 for Windows](#) (79 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

- A [pre-release](#) version for the forthcoming R-4.4.0 is available.
- 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](#).

b) Instalar RStudio:

Ingresa a: <https://posit.co/download/rstudio-desktop/>

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 215.66 MB | [SHA-256: D3C03C42](#) | Version:
2023.12.1+402 | Released: 2024-01-29

Principales conceptos de programación en R

Biblioteca (library): es el directorio donde se almacenan los paquetes se llama biblioteca. Las bibliotecas contienen las funciones predefinidas proporcionadas por el lenguaje de programación.

CRAN R: es un acrónimo de Comprehensive R Archive Network para el lenguaje de programación R. una red en la que se archivan todas las versiones de R base, así como todos los paquetes para R que han pasado por un proceso de revisión riguroso, realizado por el CRAN Team, que se encarga de asegurar su correcto funcionamiento.

CRAN mirror: mirror (espejo) es un sitio para la distribución de software. Generalmente los “mirrors” tienen todas las copias de software (bibliotecas, fuentes, códigos, funciones, etc.). Estas copias están alojadas en servidores en varios países del mundo, desde donde se pueden descargar.

Entorno de programación: es un programa o conjunto de programas que engloban todas las tareas necesarias para el desarrollo de un programa o aplicación. En la actualidad existen diferentes lenguajes de programación orientada a objetos, como C++, Objective C, Java, Ruby, Visual Basic, Visual C Sharp, Simula, Perl, TypeScript, Smalltalk, R, PHP o **Python**. C++ y Java son los dos lenguajes de programación orientada a objetos más usados

Entorno de desarrollo integrado (IDE): es una aplicación de software que ayuda a los programadores a desarrollar código de software de manera eficiente. Aumenta la productividad de los desarrolladores al combinar capacidades como editar, crear, probar y empaquetar software en una aplicación fácil de usar. Por ejemplo, el RStudio es un IDE.

Función: una función es un conjunto de instrucciones que realiza una tarea específica. Imagine una función como una pequeña máquina que toma ciertos datos (argumentos), realiza una operación y devuelve un resultado.

GitHub: es una plataforma de desarrollo colaborativo para alojar proyectos utilizando el sistema de control de versiones Git. Se utiliza principalmente para la creación de código fuente de programas computacionales. En R, por ejemplo, hay bibliotecas que no se encuentran alojadas en los CRAN mirrors y si en el GitHub personal de los docentes o investigadores.

Objetos: los objetos en programación representan cosas del mundo real, así como conceptos abstractos con sus características y comportamientos específicos. Un objeto cuenta con su estructura interna que combinan variables, funciones y estructuras de datos. Usando el nombre del objeto y la sintaxis según el lenguaje de programación, se puede visualizar los valores del objeto y llamar las funciones que tiene predefinidas. En R los principales objetos básicos son: número, carácter, vector (numérico o de caracteres), marco de datos (data frame), factor, lista, matriz, etc.

Paquete: cada paquete en R es una colección de funciones diseñadas para atender una tarea específica. Ejemplos de paquetes: agricolae, MASS, ggplot2, fBasics, etc.

4.2 Consulte los siguientes scripts base

A continuación, se presentan una serie de scripts (disponibles en el sitio del CETE) para que consulte los procedimientos básicos en el uso del entorno de programación R empleando RStudio:

R_calculadora.R

R_Objetos.R

R_importando_Excel.R

R_práctica1.R