



Tecnológico
de Monterrey

01. Presentación

Ciencia de datos para la toma de decisiones I

Jorge
Juvenal
Campos Ferreira

 juvenal.campos@tec.mx

Programa de la clase

- Presentación
- Revisión del temario
- Diagnóstico programación
- Verificación de instalación de R y Python
- Revisión detalles del curso

Sobre mí



M.C. JORGE JUVENAL CAMPOS FERREIRA.

- * **Analista de datos, México, ¿Cómo vamos? Y Fundación Novagob México**
- * **Columnista en Atiempo.TV Coahuila**

Educación Formal:



Licenciatura:

Ingeniería en Irrigación por la Universidad Autónoma Chapingo. (2009-2014).



Maestría:

Maestría en Economía por El Colegio de México (2016-2018).

Contacto:



GitHub: JuveCampos



LinkedIn: Jorge Juvenal Campos Ferreira



Twitter: @JuvenalCamposF



IG: juvenalcampos.dataviz



<https://juvenalcampos.com/>

Experiencia profesional

- **Analista de datos.**
 - CIDE - Laboratorio Nacional de Políticas Públicas.
 - México ¿Cómo vamos?
 - Fundación Novagob México
 - Clave Igualdad
- **Profesor**
 - Periodismo de datos - Maestría en Periodismo del CIDE
 - Tableros en R/Shiny - Datacrunchers
- **Periodista**
 - Columna semanal en Atiempo.TV
 - Escritor y colaborador en Nexos y Animal Político

Sobre ustedes

- ¿Cómo se llaman?
- ¿De donde son?
- ¿Por que están en esta carrera?
- ¿Que tanto saben de Ciencia de Datos?
- ¿Qué herramientas saben usar?
- ¿Qué esperan de este curso?



Instalación R y Python

- Vamos a verificar que tengan R y Python instalado.



Instalar R

<https://cran.r-project.org/>



Instalar RStudio

[https://posit-co.translate.goog/download/rstudio-desktop/?
_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc](https://posit-co.translate.goog/download/rstudio-desktop/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc)

El curso se va a dar principalmente en R y RStudio, por lo que sí es importante que verifiquen que el programa esté instalado y funcionando.

Propósito del curso

El curso busca que el estudiante:

El objetivo central de esta unidad es que el estudiante sea capaz de **impulsar transformaciones públicas de alto impacto** mediante el uso estratégico de **tecnologías digitales y disruptivas**, evaluando qué herramientas convienen para resolver problemas públicos según sus **costos, beneficios sociales y potencial de cambio**. Además, busca que pueda **diseñar propuestas de política pública resilientes** ante la incertidumbre, sustentadas en análisis **cuantitativos y cualitativos** y en la valoración comparada de alternativas. Finalmente, pretende fortalecer su capacidad para **formular preguntas e hipótesis de investigación** que permitan diagnosticar problemas políticos, económicos y sociales, y para usar la tecnología de forma **consciente y segura**, respetando la **propiedad intelectual** y construyendo una **identidad digital responsable**.

Propósito del curso

El curso busca que el estudiante:

El objetivo central de esta unidad es que el estudiante sea capaz de **impulsar transformaciones públicas de alto impacto** mediante el uso estratégico de **tecnologías digitales y disruptivas**, evaluando qué herramientas convienen para resolver problemas públicos según sus **costos, beneficios sociales y potencial de cambio**. Además, busca que pueda **diseñar propuestas de política pública resilientes** ante la incertidumbre, sustentadas en análisis **cuantitativos y cualitativos** y en la valoración comparada de alternativas. Finalmente, pretende fortalecer su capacidad para **formular preguntas e hipótesis de investigación** que permitan diagnosticar problemas políticos, económicos y sociales, y para usar la tecnología de forma **consciente y segura**, respetando la **propiedad intelectual** y construyendo una **identidad digital responsable**.

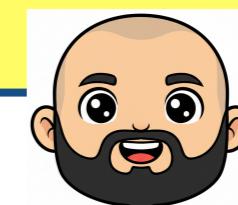
Lo que yo busco



1. Que **mejoren** sus capacidades de programación en **R**
2. Que **comprendan los fundamentos** básicos de Ciencia de Datos, Estadística y Machine Learning
3. Que **practiquen** la aplicación modelos estadísticos a los problemas que pudieran surgir

Horarios

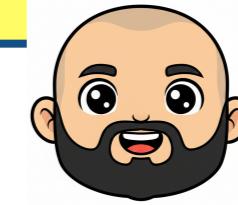
Primer tercio

Lunes	Martes (9:10-10:50)	Miércoles (9:10-10:50)	Jueves	Viernes (9:10-10:50)
				

Segundo tercio

Lunes	Martes (9:10-10:50)	Miércoles	Jueves	Viernes (9:10-10:50)
				

Tercer tercio

Lunes	Martes (9:10-10:50)	Miércoles (11:10-12:50)	Jueves	Viernes (9:10-10:50)
				

Temario y evaluación

Ver Canvas o el Github de la clase:

<https://github.com/JuveCampos/TC2001B.601-Ciencia-de-datos-ene-jun-2026>



Reto

Por definir



Reglas de clase

- Mantener el **respeto** hacia compañeros y profesores
- **Participar** de forma ordenada y respetuosa
- Evitar usar **dispositivos electrónicos** para fines ajenos a la clase
- **Traer laptop**, libreta y bolígrafos de colores
- Usar modelos de IA **cuando el profesor lo especifique**

Examen sorpresa

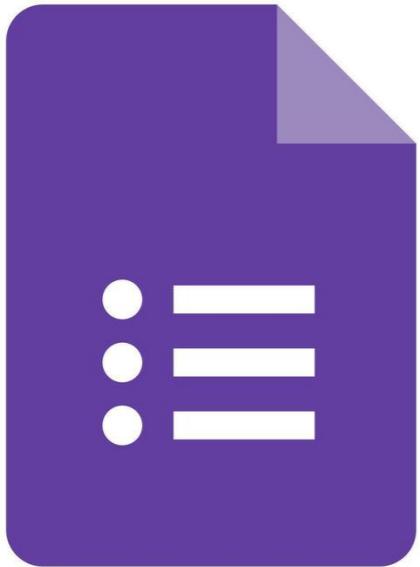
- Mecanismo para evaluar que tanto van aprendiendo.
- Sólo se aplicará en caso de que haya una alta **incidencia de uso de IA** entre los estudiantes.



Diagnostico uso herramientas Ciencia de Datos

Test diagnóstico

Ir al siguiente enlace de Google Forms:



[https://docs.google.com/forms/d/e/
1FAIpQLSdnw6ZKJPVEwwRYOjp50JRZrFqQauXf3RReiJeZqWOEQYa16A/viewform?
usp=sharing&oid=103051680334316606510](https://docs.google.com/forms/d/e/1FAIpQLSdnw6ZKJPVEwwRYOjp50JRZrFqQauXf3RReiJeZqWOEQYa16A/viewform?usp=sharing&oid=103051680334316606510)

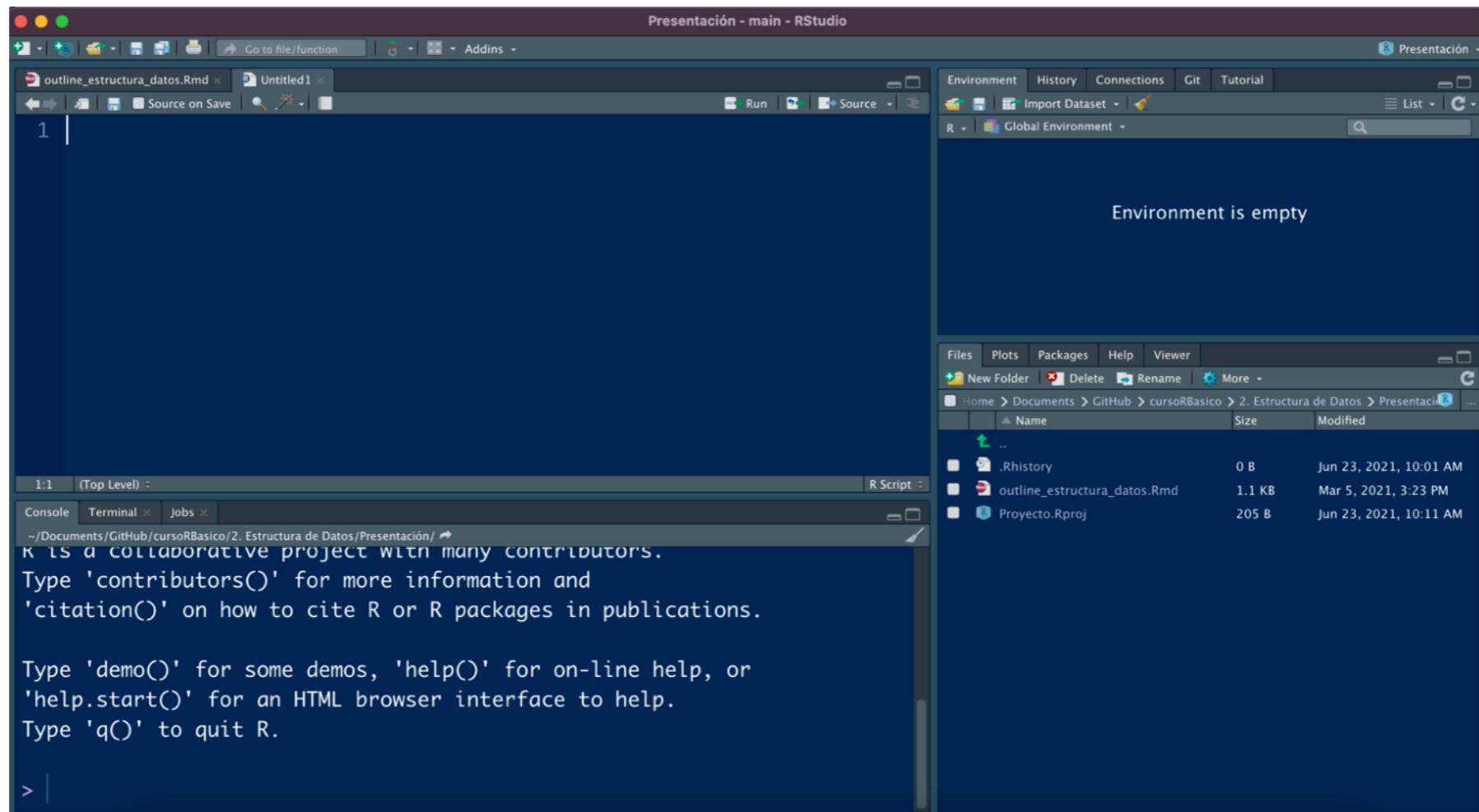
Instalar R y RStudio

Instalar R: <https://www.r-project.org/>

Instalar RStudio: <https://posit.co/download/rstudio-desktop/>



RStudio es un programa que provee un entorno de desarrollo (IDE) que nos da las herramientas necesarias para poder programar en R.



Ventanas



The screenshot illustrates the RStudio interface with several windows highlighted:

- Editor de texto**: The top-left window shows R code for data manipulation, with the first 10 lines highlighted by a yellow box.
- Consola**: The bottom-left window shows the output of the R code, with the last few lines highlighted by a green box.
- Visualizador**: The right side of the interface contains two panes:
 - Ambiente**: A pink-highlighted pane showing the Global Environment with variables like `años`, `pers...`, and `sabe...`.
 - Presentación**: A red-highlighted pane showing the Project Structure with files like `.Rhistory`, `outline_estructura_datos.Rmd`, and `Proyecto.Rproj`.

Ventanas



Editor de texto

Sección del programa en la cual registramos las instrucciones que se van a correr en R. Estas instrucciones se guardan en scripts para volver a ellos más adelante.

Acá se pueden escribir códigos de R, HTML, Python, CSS, Markdown, etc.

The screenshot illustrates the RStudio interface with several colored boxes highlighting different components:

- Editor de texto**: The code editor pane, highlighted in yellow, containing R code for data manipulation and analysis.
- Consola**: The console pane, highlighted in green, showing the output of the R code run in the editor.
- Ambiente**: The environment pane, highlighted in pink, displaying the current global variables and their values.
- Visualizador**: The file browser pane, highlighted in red, showing the project structure and files.

The code in the Editor pane:

```
1 # Librerias ----  
2 library(tidyverse)  
3  
4 # Bases de datos ----  
5 personas <- c("Juvenal", "María")  
6 años <- c(29, 30)  
7 escuelas <- c("Colmex", "UNAM")  
8 sexo <- factor(c("M", "F"),  
9                   levels = c("M", "F"))  
10 sabe_r <- c(TRUE, FALSE)
```

The output in the Consola pane:

```
+ library(tidyverse)  
+ personas  
[1] Juvenal María  
+ escuelas  
[1] Colmex UNAM  
+ sexo  
[1] M F  
Levels: M F  
+ sabe_r <- c(TRUE, FALSE)
```

Ventanas



Consola

Sección en la cual se ejecuta el código que vamos a escribir en el editor de texto.

Igualmente, podemos correr acá código de R que no requerimos guardar para más adelante.

The screenshot shows the RStudio interface with several windows open:

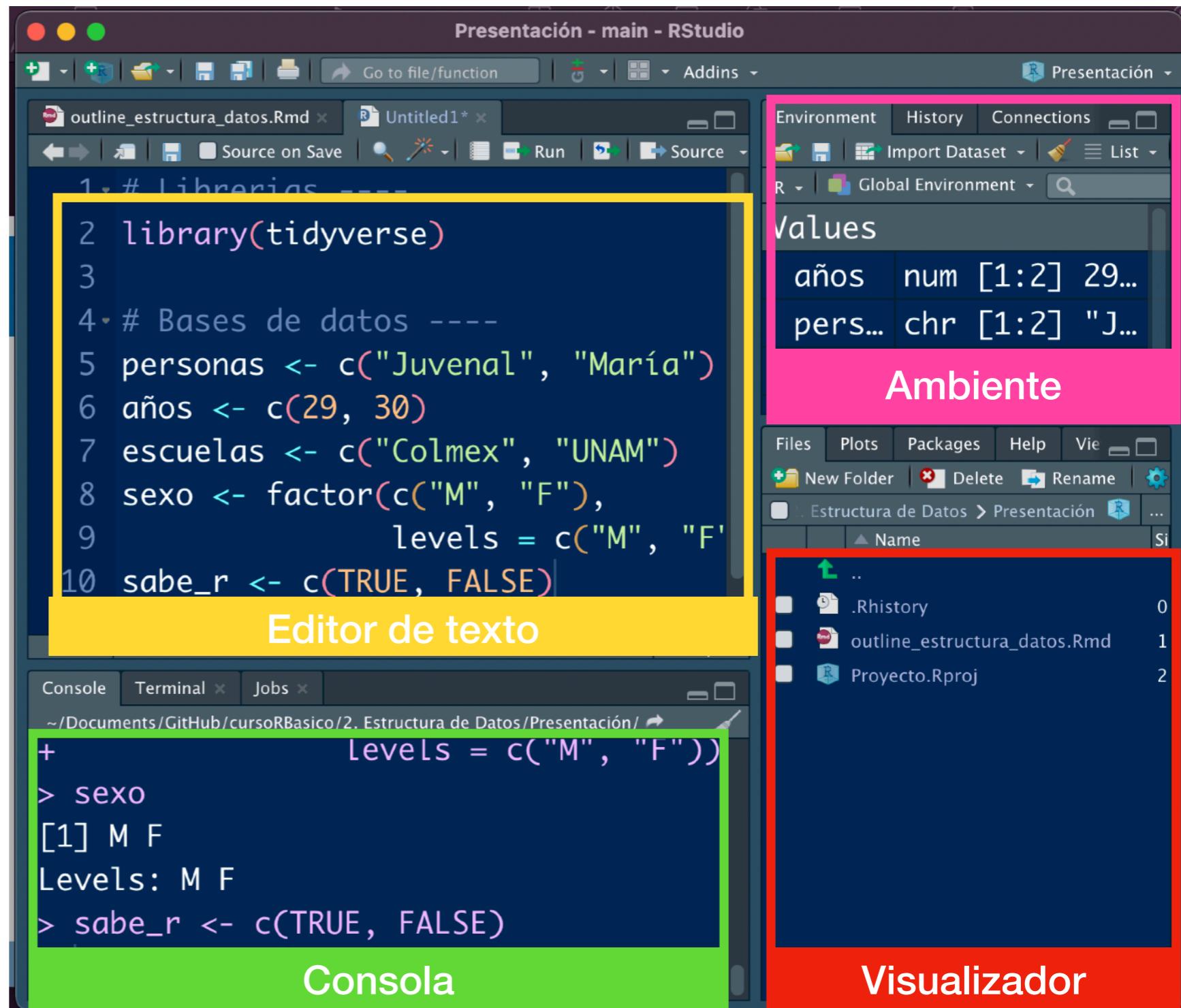
- Editor de texto**: A yellow box highlights the code editor window containing R code for creating variables (personas, años, escuelas, sexo, sabe_r) and setting levels for the sexo factor.
- Consola**: A green box highlights the console window showing the execution of the R code and its output.
- Environment**: A pink box highlights the environment pane showing the global variables: años, pers..., sabe... (with their respective types and values).
- History**: A pink box highlights the history pane showing the commandLevels = c("M", "F") and its output [1] M F.
- Connections**: A pink box highlights the connections pane.
- Global Environment**: A pink box highlights the global environment pane.
- Ambiente**: A pink box highlights the ambiente pane.
- Visualizador**: A red box highlights the visualizer pane showing the project structure: .., .Rhistory, outline_estructura_datos.Rmd, and Proyecto.Rproj.

Ventanas

Visualizador

Sección para visualizar cosas:

- 1) Archivos ubicados en nuestro directorio de trabajo
- 2) Gráficas estáticas generadas con ggplot2 o RBase.
- 3) Las librerías instaladas en nuestro RStudio.
- 4) Las visualizaciones web generadas con R.



Personalización



Para personalizar RStudio
vamos a
Tools > Global Options

Podemos configurar:

- como se visualiza el código,
- los colores de las ventanas,
- el tamaño y fuente de las letras,
- el espacio a ocupar del código,
- las cuentas para publicar resultados, etc.

The screenshot shows the 'Global Options' dialog box in R Studio. The left sidebar lists various categories: General (selected), Code, Console, Appearance, Pane Layout, Packages, R Markdown, Sweave, Spelling, Git/SVN, Publishing, Terminal, Accessibility, and Python. The main panel has tabs for Basic, Graphics, and Advanced, with Basic selected. Under the R Sessions section, the default working directory is set to ~/Library/Mobile Documents/com~apple~CloudDocs, with a 'Browse...' button. Under Workspace, there are checkboxes for restoring .RData files and saving workspaces. The History section includes options for saving history and removing duplicates. The Other section contains checkboxes for wrap-around navigation, update notifications, and crash reporting. At the bottom are OK, Cancel, and Apply buttons.

Archivos nativos de R



Al trabajar con R y RStudio, generamos cuatro tipos de archivos nativos de R, estos son los siguientes:

- *.Rproj,
- *.RData,
- *.R y
- *.rds.

Los scripts son aquellos archivos que terminan en *.r