

# Ayudantía 1

## Manipulación de objetos en R

Santiago Ortúzar

saortuza@uc.cl

Estadística Descriptiva en R

Prof. Alejandro González

Universidad Alberto Hurtado - FEN

Viernes 1 de abril de 2022

# Presentación

- Contacto: saortuza@uc.cl (mail UAH pendiente).
- Confianza:
  - Por favor, escribir SIEMPRE en caso de dudas.
  - Exigir feedback y apoyo.
- Ayudantías no son obligatorias ni tienen nota asociada.
  - Las/los invito a participar y tratar de ajustarse a la metodología que seguiremos en las ayudantías.

# Calendario

- *Tutorial instalación R y RStudio: Viernes 1 de abril (cápsula).*
- Ayudantía 1: Viernes 1 de abril (cápsula).

*Calendario todavía sujeto a confirmación por fechas de evaluaciones.*

# Plan para las ayudantías

- Primeras ayudantías: lo básico.
  - ¿Qué es R?
  - Instalar el software (R y RStudio).
  - Aprender a usar el software.

# Metodología

- 5 cápsulas y 4 clases sincrónicas.
- Todas las ayudantías irán acompañadas de una presentación y una hoja de sintaxis en R.
- No desgastarse tomando apuntes, toda la información estará en las presentaciones.

# Plan para esta sesión

- Manipulación de objetos en R:
  - Vectores.
  - Matrices.
  - Data frames.
  - Obtener información sobre variables en un data frame.

## Usando RStudio

# Observaciones

- El símbolo `>` en la consola indica que R no está corriendo ningún análisis actualmente y que podemos escribir e ingresar código.
- En principio, R lee líneas de código separadas por puntos aparte como funciones diferentes.
- Podemos correr funciones seleccionando el pedazo de texto (código) que nos interesa y apretando **Control + Enter** (*para Windows*) y **Comando + Enter** (*MacOS*).
- R permite agregar texto a nuestra hoja de sintaxis sin interpretarlo como código, siempre que esté escrito después de `#`.



# Observaciones

```
55 + 45 # "55+45" es código de R, pero esto no
```

```
## [1] 100
```

```
# Esta función es útil para comentar nuestro código
```

```
# Ayuda a quien lo lea a entenderlo mejor
```

```
# (muchas veces nosotros mismos)
```

# Los objetos en R

- Recordemos que R es un lenguaje de programación orientada a objetos.
- Un objeto en R es cualquier cosa a la que podamos asignar información numérica o alfabética.
- Para asignar información a un objeto, usaremos un símbolo de flecha hacia la izquierda: `<-`.

# Los objetos en R

```
objeto1 <- 1
```

```
objeto2 <- 1 + 1
```

```
objeto3 <- 2 - 1
```

- Estas son funciones para *crear* objetos, pero por sí mismos no nos “muestran” el objeto creado.
- Para “llamar” al objeto creado debemos escribir el nombre del objeto respectivo en la consola.

# Los objetos en R

```
objeto1
```

```
## [1] 1
```

- El objeto que aquí llamamos `objeto1` contiene el valor numérico 1.

# Los objetos en R

```
objeto2
```

```
## [1] 2
```

- objeto2 contiene la suma de  $1 + 1$ , por lo tanto al “llamarlo” nos entrega el valor numérico 2.

# Los objetos en R

```
objeto3
```

```
## [1] 1
```

- objeto3 contiene la resta de  $2 - 1$ , por lo que nos reporta un valor número 1.
- Así, si bien objeto1 y objeto3 nos entregan la misma información (el valor 1), se generan a partir de funciones diferentes.

# Los objetos en R

- Noten como cada objeto fue generado por código escrito en la **sintaxis**, que R reporta en la **consola** y se **almacena** en el environment.
- Los objetos pueden contener información no sólo numérica sino alfabética.
- Para almacenar texto en un objeto, debemos usar comillas: .

# Los objetos en R

```
objeto4 <- "texto"
```

```
objeto4
```

```
## [1] "texto"
```

```
objeto5 <- "abcd"
```

```
objeto5
```

```
## [1] "abcd"
```



# Tipos de objetos en R

- Ahora bien, en R hay muchos *tipos* de objetos, y por ahora nos interesan principalmente 2: los *vectores* y las *bases de datos (data frames)*, aunque mencionaremos otros.

# Tipos de objetos en R: vectores

- Un vector es un tipo de objeto unidimensional que contiene información alfabética o numérica.
- Cuando hablamos de vectores, hablamos de objetos que ya no contienen sólo **un** valor sino **varios** valores.

## Tipos de objetos en R: vectores

```
vector1 <- c(1, 2, 3, 4, 5)
```

```
vector1
```

```
## [1] 1 2 3 4 5
```

```
vector2 <- c("a", "b", "c", "d", "e")
```

```
vector2
```

```
## [1] "a" "b" "c" "d" "e"
```

- Como vemos, `vector1` contiene 5 valores numéricos, mientras que `vector2` contiene 5 valores alfabéticos.

## Tipos de objetos en R: vectores

- El comando `c()`, llamado *concatenate* (“concatenar”), será fundamental para poder asignar múltiples valores (esto es, más de un ítem de información) a un mismo objeto.
- Al ir creando objetos de este tipo, R nos permite avanzar de a poco hacia operaciones más complejas.
- Por ejemplo, creemos otro vector llamado `vector3` y sumémoslo a `vector1`.

## Tipos de objetos en R: vectores

```
vector3 <- c(6, 7, 8, 9, 10)
```

```
vector3
```

```
## [1] 6 7 8 9 10
```

```
vector1 + vector3
```

```
## [1] 7 9 11 13 15
```

- Al sumar `vector1 + vector3`, obtenemos una nueva información en la consola: los valores 7 9 11 13 15.
- De hecho, podríamos repetir esta operación pero almacenarla en un nuevo vector: `vector4`.

# Tipos de objetos en R: vectores

```
vector1 + vector3
```

```
## [1] 7 9 11 13 15
```

```
vector4 <- vector1 + vector3
```

```
vector4
```

```
## [1] 7 9 11 13 15
```

# Tipos de objetos en R: vectores

- Sin embargo, no podríamos hacer algo similar con `vector2`, ya que éste contiene valores alfabéticos y no numéricos.

## Tipos de objetos en R: vectores

Si ingresan el código

```
vector1 + vector2
```

R entregará el siguiente mensaje:

```
Error in vector1 + vector2 : non
```



# Tipos de objetos en R: vectores

- La razón por la que esto ocurre es que `vector1` y `vector2` pertenecen a “clases” diferentes.
- Con el comando `class()` podemos ver a qué clase corresponde cada uno.

# Tipos de objetos en R: vectores

```
class(vector1)
```

```
## [1] "numeric"
```

```
class(vector2)
```

```
## [1] "character"
```

- R indica que `vector1` es de clase numérica, mientras que `vector2` es un carácter. También hay una tercera clase de vector, “factor”, que veremos más adelante y es muy útil.
  - Dato importante: si ingresamos números a un vector, pero los números van con comillas, R los entenderá como *carácter*, no como un valor *numérico*.

# Tipos de objetos en R: vectores

```
vector5 <- c("1", "2", "3", "4", "5")
```

```
vector5
```

```
## [1] "1" "2" "3" "4" "5"
```

```
class(vector5)
```

```
## [1] "character"
```

```
str(vector5)
```

```
## chr [1:5] "1" "2" "3" "4" "5"
```

## Tipos de objetos en R: vectores

- La función `str` (structure) también reporta la clase del objeto, además de un resumen de su información.

```
str(vector5)
```

```
## chr [1:5] "1" "2" "3" "4" "5"
```

- En la práctica la gran mayoría de los análisis estadísticos que realizaremos usan vectores numéricos y factores.
- Las funciones de manipulación de texto y caracteres pueden ser muy útiles cuando tenemos que limpiar y depurar bases de datos.

# Ejercicio

- Cree un vector que contenga los valores 5, 8, 3, 2 y 10, en ese orden. Llame al vector `ejercicio`.
- Realice las siguientes operaciones con el vector:
  - Sume 9 a `ejercicio`.
  - Multiplique `ejercicio` por 2.
  - Multiplique `ejercicio` por  $5/7$ .
  - Eleve `ejercicio` al cuadrado.

# Ejercicio

```
ejercicio <- c(5, 8, 3, 2, 10)
```

```
ejercicio + 9
```

```
## [1] 14 17 12 11 19
```

```
ejercicio * 2
```

```
## [1] 10 16 6 4 20
```

```
5 / 7 * ejercicio
```

```
## [1] 3.571429 5.714286 2.142857 1.428571 7.142857
```

```
ejercicio * ejercicio # o también:
```

```
## [1] 25 64 9 4 100
```

```
ejercicio^2
```

```
## [1] 25 64 9 4 100
```

## Tipos de objetos en R: matrices

- Una matriz es un tipo de objeto bidimensional que, al igual que un vector, puede tomar valores alfabéticos o numéricos.

```
matriz1 <- matrix(1:25, nrow = 5, ncol = 5)
```

```
matriz1
```

```
##      [,1] [,2] [,3] [,4] [,5]  
## [1,]    1    6   11   16   21  
## [2,]    2    7   12   17   22  
## [3,]    3    8   13   18   23  
## [4,]    4    9   14   19   24  
## [5,]    5   10   15   20   25
```

- El comando le dice a R que cree una matriz cuyos valores vayan del 1 al 25, con 5 filas y 5 columnas.

## Tipos de objetos en R: matrices

- Podemos realizar todas las operaciones del álgebra de matrices:

```
matriz1 + matriz1
```

```
##           [,1] [,2] [,3] [,4] [,5]
## [1,]         2   12   22   32   42
## [2,]         4   14   24   34   44
## [3,]         6   16   26   36   46
## [4,]         8   18   28   38   48
## [5,]        10   20   30   40   50
```



## Tipos de objetos en R: matrices

- Nosotros no usaremos mucho las matrices propiamente tales, aunque a veces se pueden usar de forma auxiliar para asistir a otras funciones.
- Sí nos interesa poner atención a la notación de R: cuando vemos números separados por comas entre corchetes [ , ] estamos hablando de *filas* y *columnas* (en ese orden).
- Así, por ejemplo, podemos ver información sólo para filas y columnas de nuestro interés:

## Tipos de objetos en R: matrices

- Cuarta fila y quinta columna de la matriz 1:

```
matriz1[4, 5]
```

```
## [1] 24
```

```
matriz1
```

```
##      [,1] [,2] [,3] [,4] [,5]  
## [1,]    1    6   11   16   21  
## [2,]    2    7   12   17   22  
## [3,]    3    8   13   18   23  
## [4,]    4    9   14   19   24  
## [5,]    5   10   15   20   25
```

## Tipos de objetos en R: bases de datos (data frames)

- Las bases de datos consisten en conjuntos de vectores portadores de valores numéricos o alfabéticos.
- Veamos una muestra de 373 casos elegidos al azar de la III Encuesta Nacional de Participación y Consumo Cultural (2009).

# Tipos de objetos en R: bases de datos (data frames)

- La base de datos contiene las siguientes variables:
  - folio (identificador).
  - SEL\_Sexo (quizá versiones más actualizadas preguntan por género).
  - SEL\_Edad (edad como valor numérico).
  - SEL\_NEduc (nivel educativo: grado más alto alcanzado).
  - libros\_dummy: *m6\_64: Con excepción de textos escolares y considerando libros usados o nuevos, ¿ha leído usted algún libro en los últimos 12 meses?* (variable dicotómica)
  - libros\_numero: *m6\_66: En los últimos 12 meses, ¿cuántos libros ha leído?* (variable numérica)
  - libros\_frecuencia: *m6\_67: Sin considerar textos escolares, ¿con qué frecuencia acostumbra leer libros?* (variable categórica con 4 categorías)

# Tipos de objetos en R: bases de datos (data frames)

- Vamos a seleccionar el directorio donde trabajar:
  - ¿Cómo encontrar rápido el directorio?

```
getwd()
```

```
## [1] "C:/Github/uah_ayudantia_estadistica_2022/ayudantias_estadistica_descriptiva/ayudantia01"
```

- Fijar directorio:

```
setwd("C:/GitHub/uah_ayudantia_estadistica_2022/ayudantias_estadistica_descriptiva/ayudantia01")
```

# Tipos de objetos en R: bases de datos (data frames)

```
load("cc1.RData")
```

- Decirle a R dónde buscar el dataframe en mi computador.

# Tipos de objetos en R: bases de datos (data frames)

- Comandos que servirán para examinar bases de datos:  
`head()`, `tail()`, `str()`, `dim()`.

```
head(cc1a)
```

```
##      folio SEL_Sexo SEL_Edad  SEL_NEduc libros_dummy libros_numero
## 2335  8577  Hombre    27 Secundaria          No             0
## 1853  4924  Hombre    47 Secundaria          No             0
## 3311  8839   Mujer    44 Secundaria          Si              4
## 1770  8709   Mujer    34 Secundaria          No             0
## 3521  2962  Hombre    24  Superior          Si              5
## 5686  2362   Mujer    70  Primaria          Si              1
##
##               libros_frecuencia
## 2335      Menos de una vez al año
## 1853      Menos de una vez al año
## 3311 1 día por semana - 1 día al mes
## 1770      Menos de una vez al año
## 3521 1 día por semana - 1 día al mes
## 5686      Menos de una vez al mes
```

# Tipos de objetos en R: bases de datos (data frames)

```
tail(cc1a)
```

```
##      folio SEL_Sexo SEL_Edad SEL_NEduc libros_dummy libros_numero
## 7631  4310  Hombre     37 Secundaria      Sí             1
## 4977  7062   Mujer     48  Primaria      No             0
## 1285  7734   Mujer     61  Superior      Sí             5
## 6407  2756  Hombre     55 Secundaria      No             0
## 5262  2994   Mujer     50  Primaria      Sí             1
## 4333  8578  Hombre     28  Superior      No             0
##
##               libros_frecuencia
## 7631 1 día por semana - 1 día al mes
## 4977      Menos de una vez al año
## 1285      Todos o casi todos los días
## 6407      Menos de una vez al año
## 5262 1 día por semana - 1 día al mes
## 4333      Menos de una vez al año
```



# Tipos de objetos en R: bases de datos (data frames)

```
str(cc1a)
```

```
## 'data.frame':    373 obs. of  7 variables:
## $ folio          : int  8577 4924 8839 8709 2962 2362 6247 3396 4163 6779
## $ SEL_Sexo       : chr  "Hombre" "Hombre" "Mujer" "Mujer" ...
## $ SEL_Edad       : int  27 47 44 34 24 70 25 49 49 68 ...
## $ SEL_NEduc      : Factor w/ 3 levels "Primaria","Secundaria",...: 2 2 2 2
## $ libros_dummy   : chr  "No" "No" "Sí" "No" ...
## $ libros_numero  : int  0 0 4 0 5 1 0 2 2 2 ...
## $ libros_frecuencia: Factor w/ 4 levels "Todos o casi todos los días",...: 4
## - attr(*, "na.action")= 'omit' Named int [1:27] 9 25 29 44 74 76 119 132 14
## ..- attr(*, "names")= chr [1:27] "7054" "355" "2268" "2993" ...
```

# Tipos de objetos en R: bases de datos (data frames)

```
dim(cc1a)
```

```
## [1] 373 7
```

# Tipos de objetos en R: bases de datos (data frames)

- Para acceder a una variable en particular dentro de la base de datos, debemos usar el comando \$.
- Ejemplos de descriptivos univariados:
  - Conteo de frecuencias de sexo y lectura de libros en el último año:

```
table(cc1a$SEL_Sexo)
```

```
##  
## Hombre  Mujer  
##      172    201
```

```
table(cc1a$libros_dummy)
```

```
##  
## No  Sí  
## 198 175
```

# Tipos de objetos en R: bases de datos (data frames)

- Para acceder a una variable en particular dentro de la base de datos, debemos usar el comando \$.
- Ejemplos de descriptivos univariados:
  - Proporciones de cada categoría de sexo y lectura en los últimos 12 meses:

```
prop.table(table(cc1a$SEL_Sexo))
```

```
##  
##      Hombre      Mujer  
## 0.461126 0.538874
```

```
prop.table(table(cc1a$libros_dummy))
```

```
##  
##           No           Sí  
## 0.5308311 0.4691689
```

# Tipos de objetos en R: bases de datos (data frames)

- Para acceder a una variable en particular dentro de la base de datos, debemos usar el comando \$.
- Ejemplos de descriptivos univariados:
  - Frecuencias de qué tan seguido leen las personas encuestadas:

```
table(cc1a$libros_frecuencia)
```

```
##
##      Todos o casi todos los días 1 día por semana - 1 día al mes
##                                58                                75
##      Menos de una vez al mes      Menos de una vez al año
##                                42                                198
```

# Tipos de objetos en R: bases de datos (data frames)

- Para acceder a una variable en particular dentro de la base de datos, debemos usar el comando \$.
- Ejemplos de descriptivos univariados:
  - Proporciones de qué tan seguido leen las personas encuestadas:

```
prop.table(table(cc1a$libros_frecuencia))
```

```
##  
##      Todos o casi todos los días 1 día por semana - 1 día al mes  
##                0.1554960                0.2010724  
##      Menos de una vez al mes      Menos de una vez al año  
##                0.1126005                0.5308311
```

# Tipos de objetos en R: bases de datos (data frames)

- Proporciones de qué tan seguido leen las personas encuestadas:
- Para mejor presentación de las proporciones:

```
100 * prop.table(table(cc1a$libros_frecuencia))
```

```
##  
##      Todos o casi todos los días 1 día por semana - 1 día al mes  
##              15.54960                      20.10724  
##      Menos de una vez al mes      Menos de una vez al año  
##              11.26005                      53.08311
```

# Tipos de objetos en R: bases de datos (data frames)

- Proporciones de qué tan seguido leen las personas encuestadas:
- Para mejor presentación de las proporciones:

```
round(100 * prop.table(table(cc1a$libros_frecuencia)), digits = 1)
```

```
##  
##      Todos o casi todos los días 1 día por semana - 1 día al mes  
##                      15.5                      20.1  
##      Menos de una vez al mes      Menos de una vez al año  
##                      11.3                      53.1
```



# Tipos de objetos en R: bases de datos (data frames)

- Para acceder a una variable en particular dentro de la base de datos, debemos usar el comando \$.
- Ejemplos:
  - ¿Cuáles son la media y la mediana de libros leídos el último año?

```
mean(cc1a$libros_numero)
```

```
## [1] 1.865952
```

```
median(cc1a$libros_numero)
```

```
## [1] 0
```

# Tipos de objetos en R: bases de datos (data frames)

- Al igual que con las matrices, podemos utilizar los corchetes para señalar filas y columnas de nuestro interés.
- Por ejemplo, veamos qué valor tiene en la fila 100 la pregunta sobre frecuencia de lectura:

```
cc1a[100, "libros_frecuencia"]
```

```
## [1] Todos o casi todos los días
```

```
## 4 Levels: Todos o casi todos los días ... Menos de una vez al año
```

```
# o lo mismo escrito de otra forma:
```

```
cc1a[100, 7]
```

```
## [1] Todos o casi todos los días
```

```
## 4 Levels: Todos o casi todos los días ... Menos de una vez al año
```

```
# porque libros_frecuencia es la séptima variable (columna)  
# de nuestro data frame
```

# Tipos de objetos en R: bases de datos (data frames)

- O los valores de las tres variables sobre lectura en las filas 150 y 251:

```
cc1a[150, c("libros_dummy", "libros_numero", "libros_frecuencia")]
```

```
##      libros_dummy libros_numero      libros_frecuencia  
## 8049           No           0 Menos de una vez al año
```

```
cc1a[251, c("libros_dummy", "libros_numero", "libros_frecuencia")]
```

```
##      libros_dummy libros_numero      libros_frecuencia  
## 536             Sí           2 Todos o casi todos los días
```

*# o lo mismo escrito de otra forma:*

```
cc1a[251, 5:7]
```

```
##      libros_dummy libros_numero      libros_frecuencia  
## 536             Sí           2 Todos o casi todos los días
```

*# porque las variables son la número 5, 6 y 7  
# de nuestro data frame*

# Tipos de objetos en R: bases de datos (data frames)

- Veamos ahora algunos descriptivos bivariados (variable numérica según variable categórica):
  - Cómo se distribuyen las *medias* de una variable numérica según las diferentes *categorías* de nivel educativo y sexo.

```
tapply(cc1a$libros_numero, cc1a$SEL_NEduc, mean)
```

```
##   Primaria Secundaria   Superior  
##   1.021505   1.804878   3.080000
```

*# media de libros leídos por año según nivel educativo*

```
tapply(cc1a$libros_numero, cc1a$SEL_Sexo, mean)
```

```
##   Hombre   Mujer  
## 1.110465 2.512438
```

*# media de libros leídos por año según sexo*

# Tipos de objetos en R: bases de datos (data frames)

- Veamos ahora algunos descriptivos bivariados (asociación entre variables categóricas, que se registra en una tabla de contingencia):

```
table(cc1a$libros_frecuencia, cc1a$SEL_NEduc)
```

```
##
##               Primaria Secundaria Superior
## Todos o casi todos los días           14           26           18
## 1 día por semana - 1 día al mes       13           51           11
## Menos de una vez al mes                4           21           17
## Menos de una vez al año                62          107           29
```

```
# frecuencia de lectura según nivel educativo
```

```
# guardemos esto como objeto para mayor simplicidad:
```

```
lectura_NEduc <- table(cc1a$libros_frecuencia, cc1a$SEL_NEduc)
```

```
# lo mismo de otra manera:
```

```
xtabs(~ libros_frecuencia + SEL_NEduc, data = cc1a)
```

```
##
## libros_frecuencia      SEL_NEduc
##               Primaria Secundaria Superior
## Todos o casi todos los días           14           26           18
## 1 día por semana - 1 día al mes       13           51           11
## Menos de una vez al mes                4           21           17
```

# Tipos de objetos en R: bases de datos (data frames)

- Veamos ahora algunos descriptivos bivariados (asociación entre variables categóricas):
  - Se entiende mejor con proporciones:

```
round(100 * prop.table(lectura_NEduc), 2)
```

```
##  
##  
##          Primaria Secundaria Superior  
## Todos o casi todos los días          3.75          6.97          4.83  
## 1 día por semana - 1 día al mes      3.49         13.67          2.95  
## Menos de una vez al mes              1.07          5.63          4.56  
## Menos de una vez al año              16.62         28.69          7.77
```

# Tipos de objetos en R: bases de datos (data frames)

- Veamos ahora algunos descriptivos bivariados (asociación entre variables categóricas):
  - Se entiende mejor con proporciones (*porcentaje de columna*):

```
round(100 * prop.table(lectura_NEduc, margin = 2), 2)
```

```
##
##
##      Primaria Secundaria Superior
## Todos o casi todos los días      15.05      12.68      24.00
## 1 día por semana - 1 día al mes  13.98      24.88      14.67
## Menos de una vez al mes          4.30      10.24      22.67
## Menos de una vez al año          66.67      52.20      38.67
```

```
# el porcentaje de columna se especifica con margin=2
```

# Tipos de objetos en R: bases de datos (data frames)

```
##
##
##      Primaria Secundaria Superior
## Todos o casi todos los días      15.05      12.68      24.00
## 1 día por semana - 1 día al mes  13.98      24.88      14.67
## Menos de una vez al mes          4.30      10.24      22.67
## Menos de una vez al año          66.67      52.20      38.67
```

- ¿Cómo interpretar el porcentaje de columna? De arriba hacia abajo, todas las columnas suman 100 % (no así las filas de izquierda a derecha).
- Nos indican cómo varía la frecuencia de lectura según nivel educativo.
- Así, entre quienes están/alcanzaron la primaria, alrededor del 67 % declara que lee menos de un libro al año, mientras que el 15 % lee todos o casi todos los días.
- Entre quienes están/alcanzaron la educación superior, alrededor del 39 % no lee casi nunca, mientras que el 24 % lee todos o casi todos los días.



# Tipos de objetos en R: bases de datos (data frames)

## ■ Ahora con *porcentaje de fila*:

```
round(100 * prop.table(lectura_NEduc, margin = 1), 2)
```

```
##
##
##      Primaria Secundaria Superior
## Todos o casi todos los días      24.14      44.83      31.03
## 1 día por semana - 1 día al mes  17.33      68.00      14.67
## Menos de una vez al mes          9.52      50.00      40.48
## Menos de una vez al año          31.31      54.04      14.65
```

```
# el porcentaje de columna se especifica con margin=1
```

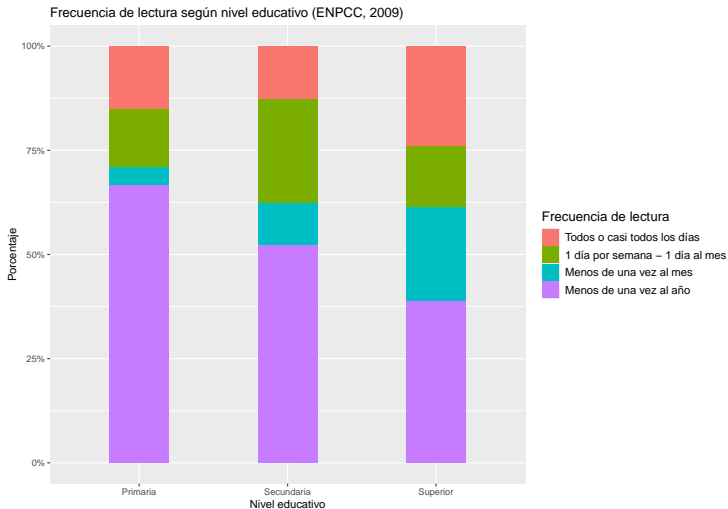
- ¿Cómo interpretar el porcentaje de fila? De izquierda a derecha, todas las filas suman 100 % (no así las columnas de arriba hacia abajo).
- Nos indica cómo varía el nivel educativo en cada tramo/grupo/categoría de frecuencia de lectura.
- Así, entre quienes leen todos/casi todos los días, alrededor del 24 % alcanzó la educación primaria, mientras que el 31 % alcanzó la educación superior.
- Entre quienes leen menos de un libro al año, el 31 % alcanzó la primaria, mientras que el 15 % alcanzó la educación superior.

## Propuesta de ejercicio

## Propuesta de ejercicio (totalmente voluntario)

- Utilizando la base de datos `cc1a` (archivo `cc1.Rdata`) que analizamos esta clase, y el código visto en clases y en ayudantía, haga un análisis de la lectura en Chile según nuestros datos.
- Puede incluir tablas de descriptivos univariados o bivariados, además de gráficos.

# Propuesta de ejercicio (totalmente voluntario)



## Propuesta de ejercicio (totalmente voluntario)

- Más allá del código, ¿qué *información* o *aprendizaje* nos dejan los datos? ¿Hay alguna hipótesis que le parezca pertinente investigar? Sea creativo/a, estamos aprendiendo y equivocarnos da lo mismo.
- Este ejercicio se trata en parte de usar el código, pero también de formular una narración que resuma lo que aprendimos de los datos.
- Deberíamos aspirar a analizar datos para poder evaluar argumentos y teorías.

## Bibliografía

## Referencias sugeridas

En castellano:

- Hernández, F., & Usuga, O. (2021a). Introducción. En Manual de R. <https://fhernanb.github.io/Manual-de-R/> (disponible gratuitamente online).
- Hernández, F., & Usuga, O. (2021b). Tipos de objetos. En Manual de R. <https://fhernanb.github.io/Manual-de-R/objetos.html> (disponible gratuitamente online).

## Referencias sugeridas

En inglés:

- Paradis, E. (2013). *R for Beginners*. 1–52.



## Referencias complementarias (consulta)

- Beaujean, A. A. (2014). Introduction to R. En *Latent Variable Modeling Using R* (pp. 1–33). Routledge.
- Gorgas, J., & Cardiel, N. (s.f.). Tema 1: Introducción al paquete de software R.  
[https://www.ucm.es/data/cont/docs/339-2016-09-29-Introduccion%20a%20R\\_v1617.pdf](https://www.ucm.es/data/cont/docs/339-2016-09-29-Introduccion%20a%20R_v1617.pdf) (infografías).
- Zuur, A. F., Ieno, E. N., & Meesters, E. H. W. G. (2009). Ch. 2: Getting Data into R. En *A Beginner's Guide to R* (pp. 29–56). Springer.

# Sobre la lectura

- La lectura es de carácter voluntario y no se evaluará.
- Se recomienda leer 1 de las lecturas sugeridas y revisar las complementarias como material de consulta.