

# Encuesta Demográfica Retrospectiva 2019: Tutorial para el procesamiento de datos N° 2

*Medalla de oro a la calidad  
en la Gestión Pública*



**IRAM-ISO 9001:2015**

# ¿Qué encontrará en este documento?

A través de este material buscamos facilitar la explotación de la primera encuesta demográfica retrospectiva de estadística pública del país: Encuesta Demográfica Retrospectiva 2019 (EDER-CABA)<sup>1</sup>.

Dado su carácter retrospectivo, la EDER-CABA cuenta con una estructura de datos diferente al modo en que se presenta la información de las encuestas transversales, requiriendo un tratamiento particular de sus bases de datos.

Este documento es una continuación del Tutorial N°1. En dicho material se presentan los alcances del uso de la EDER-CABA en cuanto a su potencial de explotación longitudinal y un paso a paso con los principales comandos de programación para descargar y preparar las bases de datos para realizar análisis descriptivo, a partir de un ejemplo de aplicación.

En este Tutorial N°2, se incluyen otros comandos desarrollados en Lenguaje R para facilitar a los usuarios la fusión de las bases de datos con estructura matriz años-persona y su procesamiento, tomando

un nuevo ejemplo de aplicación. Dichos comandos pueden ser modificados según la selección de variables que resulten de su interés.

Adicionalmente, se presenta una aproximación conceptual a las principales técnicas de análisis longitudinal que pueden aplicarse con la información retrospectiva que arroja la EDER-CABA:

- 1) tablas de vida (con casos truncados),
- 2) modelos de regresión de tiempo discreto y
- 3) análisis de secuencias.

El desarrollo de las técnicas se complementa con ejemplos que aplican dichas técnicas en base a diferentes encuestas (transversales con preguntas retrospectivas, EDER-CABA y EDER-México) y con material ampliatorio para quienes quieran profundizar en sus alcances.

<sup>1</sup> Documento elaborado por el equipo Ubacyt 20020220400008BA (dirección de Sabrina Ferraris, Instituto de Investigaciones Gino Germani, Universidad Nacional de Buenos Aires). Autoría del documento: Mariela Giacoponello, con coordinación/supervisión de Mario Martínez Salgado y Sabrina Ferraris, y colaboración de investigadores asociados al proyecto.

## Índice

<b>1. Tutorial para procesar los datos de la EDER-CABA 2019</b>	<b>3</b>
1.1 Selección de variables y fusión de bases de datos con matriz años-persona	3
1.2 Construcción de la variable de interés a partir de la fusión de las bases de datos	4
1.3 Procesamiento de comandos de estadística descriptiva	5
1.3.1 Selección de variables y fusión de bases de datos con matriz años-persona	5
1.3.2 Construcción de la variable de interés a partir de la fusión de las bases de datos	8
1.3.3 Procesamiento de comandos de estadística descriptiva	9
<b>2. ¿Qué técnicas estadísticas longitudinales se pueden aplicar con la EDER-CABA?</b>	<b>12</b>
2.1 Aproximación teórica y ejemplos de aplicación	12
2.1.1 Tabla de vida (con casos truncados)	12
2.1.2 Modelos de Regresión de Tiempo Discreto	14
2.1.3 Análisis de Secuencias	16
<b>3. Bibliografía citada</b>	<b>19</b>

# 1. Tutorial para procesar los datos de la EDER-CABA 2019

Se resumen las acciones que se realizan en el paso a paso con los comandos en Lenguaje R:

## 1.1 Selección de variables y fusión de bases de datos con matriz años-persona

Como se describe en el Tutorial N° 1, la información de la EDER-CABA 2019 se encuentra distribuida en cinco bases de datos, entre las cuales tres tienen una estructura matriz años-persona, es decir, cada fila representa los años calendario o la edad de las personas (años-persona) y cada columna refiere a los eventos y sus cambios de estado a lo largo del tiempo.

En este tutorial aplicamos los comandos para la fusión de estas tres bases de datos, tomando en cuenta que, aquellas referidas a la historia de eventos conyugal ('eder2019\_usuarios\_retro\_conyuge') y de hijo/as ('eder2019\_usuarios\_retro\_hijo'), tienen una estructura en la que cada año calendario de las personas se repite según la cantidad de cónyuges e hijo/as.

Por este motivo, cada base de datos tiene una cantidad de registros diferente, requiriendo de un tratamiento particular para la fusión.

A modo de ejemplo, construimos una variable ('Situación familiar') que combina la historia de las uniones conyugales y de la ma/paternidad, para analizar las diferencias según género y generación.

Para ello, trabajamos con las variables móviles sobre la historia conyugal y de hijo/as y las variables fijas de sexo (como proxy de género) y cohorte que se encuentran distribuidas en las tres bases de datos con matriz años-persona.

Tabla n°1. Distribución de las variables de interés en las bases con matriz años-persona de la EDER-CABA

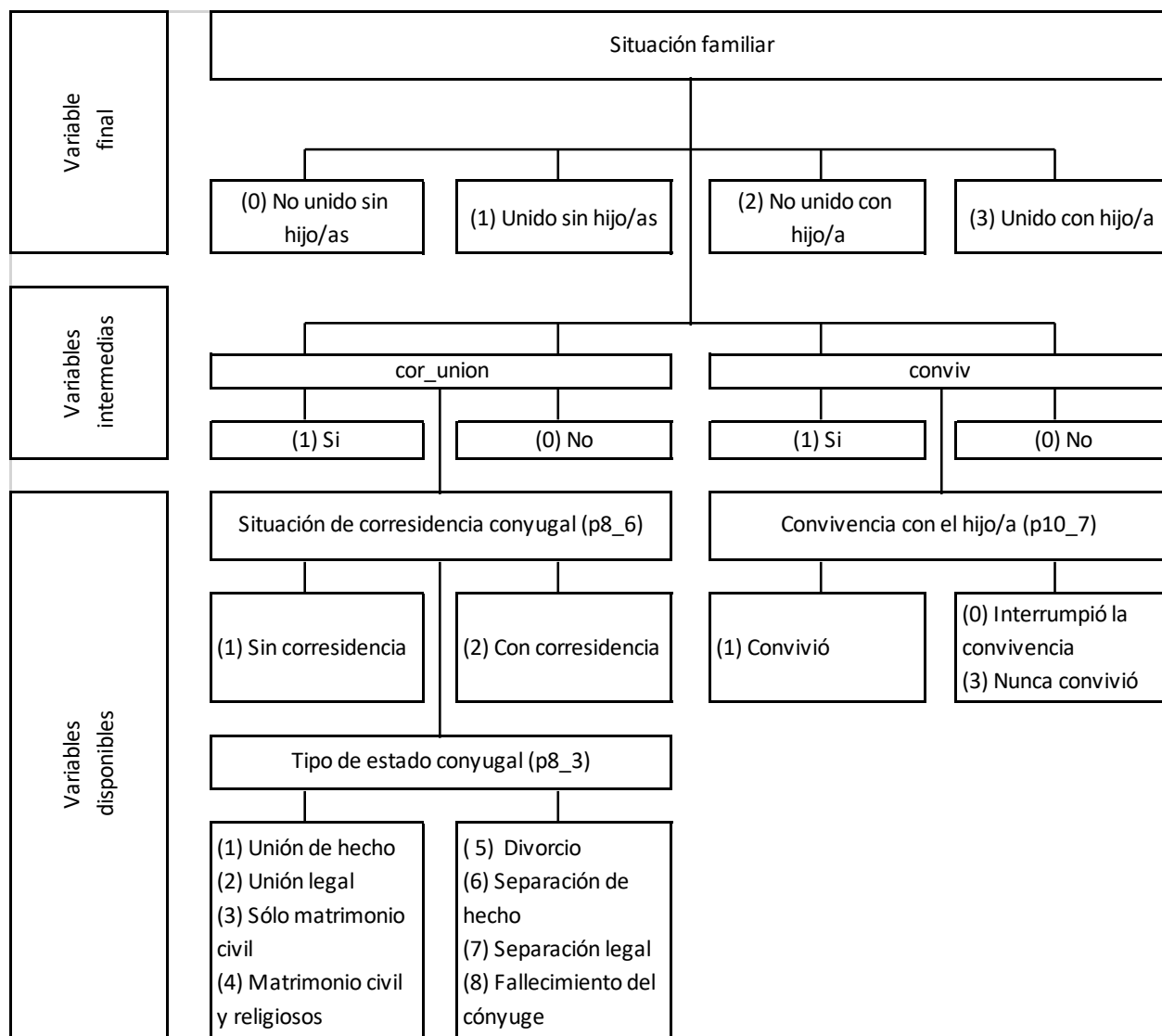
Nombre base de datos	Cantidad de registros	Variables seleccionadas		
		Código	Nombre	Categorías
eder2019_usuarios_retro	65.200	p2	Sexo	1 Varón 2 Mujer
		cohorte	Cohorte de nacimiento	1 1948-1952 2 1968-1972 3 1978-1982
eder2019_usuarios_retro_conyuge	77.524	p8_3	Tipo de estado conyugal	1 Unión de hecho 2 Unión legal 3 Sólo matrimonio civil 4 Sólo matrimonio civil y religiosos 5 Divorcio 6 Separación de hecho 7 Separación legal 8 Fallecimiento del cónyuge
		p8_6	Situación de coresidencia conyugal	1 Sin coresidencia 2 Con coresidencia
eder2019_usuarios_retro_hijo	122.405	p10_7	Convivencia con el hijo/a	0 Interrumpió la convivencia 1 Convivió 3 Nunca convivió

## 1.2 Construcción de la variable de interés a partir de la fusión de las bases de datos

La variable ‘Situación familiar’ nos permitirá identificar a cada edad de las personas las cuatro combinaciones posibles según hayan transitado o no el evento de la unión conyugal y la ma/paternidad: ‘No unido sin hijo/as’, ‘Unido sin hijo/as’, ‘No unido con hijo/as’, y ‘Unido con hijo/as’.

Previamente, elaboramos dos variables intermedias dicotómicas sobre la unión (‘cor\_union’) y la ma/paternidad (‘conviv’) que incluyen la noción de coresidencia, a modo de explotar la batería de variables disponibles en la encuesta.

Fig. nº1. Proceso de transformación de las variables de origen a la variable de interés



### 1.3 Procesamiento de comandos de estadística descriptiva

Nos preguntamos cuántas personas han transitado los dos eventos de adultez en simultáneo a una edad determinada. Interesa también conocer si hay desigualdades de género con relación a la conformación de las familias así como su variación entre generaciones.

A partir de estos datos elaboramos tablas de contingencia para conocer cómo se distribuyen los estados de la variable a una edad determinada (en el ejemplo, se toman los 35 años). Adicionalmente, incorporamos entre los cruces de variables la información sobre el origen social de las personas entrevistadas que se encuentra en la base 'eder2019\_usuarios\_IOS'.

Incluimos gráficas para poder visualizar no sólo la "foto" a un momento de las biografías, sino también la "película" sobre el comportamiento de la variable en un período de tiempo. El ejemplo incorpora la temporalidad entre los 20 y los 40 años y lo comparamos entre las generaciones para identificar los cambios en el tiempo histórico.

## Comandos para fusionar y procesar bases de datos con matriz años-persona

### Preliminares

Antes de comenzar, limpiar el área de trabajo (Environment)

```
rm(list = ls())
```

Por defecto la carpeta Documentos es el directorio de trabajo. Para cambiarlo debemos seguir los siguientes pasos:

1. En el menú principal pulsar la opción **Session**
2. Desplegar las opciones de **Set Working Directory**
3. Elegimos la opción **Choose Directory...**
4. En la ventana emergente seleccionar la carpeta de interés

Otra opción es utilizar la función `setwd()`

```
setwd("Ruta donde aloja las bases de datos")
```

### Descargar y fusionar las bases de datos

Los datos de la EDER-CABA se encuentran alojados en el sitio web de la Dirección General de Estadística y Censos. Al pulsar en el siguiente [enlace](#) se puede acceder a ellos. Tras descargar la carpeta comprimida *Base\_usuarios\_EDER-CABA\_2019\_con\_IOS* extraemos el contenido y lo ponemos en la carpeta de trabajo.

Cargamos las bases de datos con la función `read.table()`.

```
eder_retro <- read.table("eder2019_usuarios_
retro.txt", header =
TRUE, sep = "\t",
fileEncoding = "UCS-2LE")

eder_hijo <- read.table("eder2019_usuarios_
retro_hijo.txt",
header = TRUE, sep = "\t",
fileEncoding = "UCS-2LE")

eder_conyu <- read.table("eder2019_usuari-
os_retro_conyuge.txt",
header = TRUE, sep = "\t",
fileEncoding = "UCS-2LE")
```

#### 1.3.1 Selección de variables y fusión de bases de datos con matriz años-persona

### Selección de variables en cada base de datos

Una vez descargadas las bases, trabajaremos con aquellas variables que nos sirven para construir la variable final Situación familiar. Para ello, en `eder_conyu` modificamos la variable `p8_3` (tipo de estado conyugal, con ocho categorías) a una dicotómica que llamaremos `union`: (0) No unido, (1) Unido.

```
eder_conyu$union <- ifelse(eder_conyu$p8_3
%in% c(1:4),1,0)

table(eder_conyu$union, useNA = "always")

##
##      0      1 <NA>
## 52530 24994      0
```

Incluimos la variable sobre la coresidencia con el/ la cónyuge (variable `p8_6`) a la caracterización de la variable `union` que llamaremos `cor_union`: (0) No unido, (1) Unido con coresidencia.

```

eder_conyu$cor_union <- ifelse(eder_co-
nyu$union == 1 &
    eder_conyu$p8_6 == 2, 1, 0)

table(eder_conyu$cor_union, useNA = "always")

##
##      0      1 <NA>
## 52754 24770      0

table(eder_conyu$union, eder_conyu$cor_union,
useNA = "always")

##
##           0      1 <NA>
##  0      52530      0      0
##  1          224 24770      0
## <NA>         0      0      0

```

Una vez recategorizada la información sobre la unión en `eder_conyu`, podemos trabajar con los datos sobre los hijo/as de las personas encuestadas en `eder_hijo`.

Primero dicotomizaremos la variable móvil `p10_7` (convivencia con el hijo/a) como `conviv`: (0) No, (1) Si. Esta variable incorpora el componente de coresidencia a la tenencia de hijo/as.

```

eder_hijo$conviv <- ifelse(eder_hijo$p10_7 ==
1, 1, 0)

table(eder_hijo$conviv, useNA = "always")

##
##      0      1 <NA>
## 83936 38469      0

```

Puede que nos interese conocer la edad en que las personas de la encuesta dan comienzo a la ma/paternidad. Esto nos va a servir, por ejemplo, para identificar una edad de corte para el análisis descriptivo. Para ello, primero hay que crear una variable con la edad del hijo (`ed_hijo`) a partir de la de año de nacimiento del hijo/a (`p10_2b`).

```

eder_hijo$ed_hijo <- eder_hijo$annio_retro -
eder_hijo$p10_2b
eder_hijo$ed_hijo[eder_hijo$ed_hijo < 0] <- NA
eder_hijo$ed_hijo[eder_hijo$conviv != 1] <- NA

```

Luego, asociar la edad del hijo/a (`ed_hijo`) con la edad del ma/padre (`edad_retro`) para obtener la distribución de la edad del padre o la madre al momento del nacimiento del primogénito/a. Esto puede obtenerse a partir de una tabla o histograma. En este caso puede observarse que los casos se concentran entre los 26 y los 36 años.

```

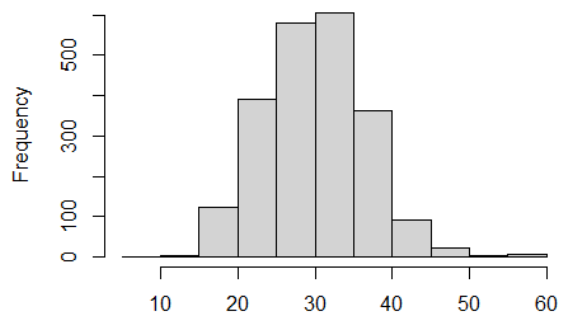
table(eder_hijo$edad_retro[eder_hijo$annio_
retro == eder_hijo$p10_2b],
    useNA = "always")

##
##      9  14  15  16  17  18  19  20
21  22  23  24  25  26  27  28
##      1   2   2   7  12  23  30  52
58  63  75  96  99 105 115 115
##     29  30  31  32  33  34  35  36
37  38  39  40  41  42  43  44
##    119 128 143 132 121 115  95 119
80  70  43  51  33  22  19   8
##    45  46  47  48  49  50  51  55
56  57  58  60 <NA>
##    10  11   3   5   3   1   3   1
2   2   1   1   0

hist(eder_hijo$edad_retro[eder_hijo$annio_
retro == eder_hijo$p10_2b])

```

**f eder\_hijo\$edad\_retro[eder\_hijo\$annio\_retro == ed**



`eder_hijo$edad_retro[eder_hijo$annio_retro == eder_hijo$p10_2`

Recordemos que la información sobre conyugalidad e hijo/as están en bases de datos diferentes. En estas bases, sólo están los casos (variable `id`) que transitaron los eventos. Además, en cada base se multiplican los años-persona (filas) por cada cónyuge o hijo/a de cada `id`. Por ejemplo, el `id = 4` tenía 70 años al momento de la encuesta y tuvo 1 cónyuge y 4 hijos.

En la base de datos `eder2019_usuarios_retro_conyuge` el `id = 4` tiene 71 filas (la cantidad de años al momento de la encuesta, recordar que incluye la edad 0, multiplicado por 1 cónyuge) y en la `base eder2019_usuarios_retro_hijo` tiene 284 filas (la cantidad de años al momento de la encuesta multiplicado por los 4 hijo/as).

En cambio, el `id = 20` no está en ninguna de las dos bases de datos porque no tuvo cónyuges ni hijo/as al momento de la encuesta.

Es importante tener esto en cuenta para comprender el proceso de fusión de bases que mostraremos a continuación.

## Fusión de las bases de datos

Una vez construidas las variables intermedias sobre unión y ma/paternidad, podemos fusionar las bases de datos. Para ello, creamos el objeto `eder` en el que incluiremos las variables fijadas que están en `eder_retro` que nos interesa cruzar: `p2` (sexo) y `cohorta` (cohorta de nacimiento). También incluiremos las variables de asociación entre bases (`id` y `edad_retro`).

```
eder <- eder_retro[, c("id", "edad_retro", "p2", "cohorta")]
```

Primero fusionaremos las variables de unión de `eder_conyu` y `eder_hijo` al objeto `eder`. Recordemos que en `eder` están todas las personas encuestadas ( $n = 1220$ ) por lo que cuando fusionemos las variables de `eder_conyu` y `eder_hijo` el sistema imputará NA (*Not Available*) para los casos que no transitaron los eventos de conyugalidad y ma/paternidad.

También imputará NA en los `id` que sí hayan transitado por los eventos de conyugalidad y ma/paternidad, pero cuya cantidad total de uniones e hijo/as sea menor al máximo de casos por `id` de la base de datos.

En el caso de la conyugalidad, hay casos con hasta 5 uniones por los que hay que repetir la fusión esa cantidad de veces para pasar las filas de cada unión como variables y quienes no tienen esa cantidad de uniones, el sistema le agrega NA. En la base de datos sobre la historia de eventos de los hijo/as hay declarados hasta 13 hijo/as por caso.

```
table(eder_conyu$conyuge, useNA = "always")
```

```
##
##      1      2      3      4      5 <NA>
## 59584 14522 2968  329  121     0
```

```
table(eder_hijo$hijo, useNA = "always")
```

```
##
##      1      2      3      4      5      6      7
##      8      9     10     11     12     13
## 51991 39527 17569 6722 3010 1367  923
## 556   256   206   138   70   70
## <NA>
##      0
```

Previo a la fusión, ordenamos la base de acuerdo con el identificador y a la variable `edad_retro`.

```
eder_retro <- eder_retro[order(eder_retro$id,
                              eder_retro$edad_retro),]
```

```
eder_conyu <- eder_conyu[order(eder_conyu$id,
                              eder_conyu$edad_retro),]
```

```
eder_hijo <- eder_hijo[order(eder_hijo$id,
                             eder_hijo$edad_retro),]
```

Fusión de `eder` con `eder_conyu`

Pegado de la información de la pareja 1.

```
eder <- merge(eder[, c("id", "edad_retro", "p2", "cohorta")],
             eder_conyu[eder_conyu$conyuge == 1,
                       c("id", "edad_retro", "cor_union")],
             by = c("id", "edad_retro"),
             all.x = TRUE)
names(eder)[c(5)] <- c("cor_union1")
```

Pegado de la información de la pareja 2.

```
eder <- merge(eder, eder_conyu[eder_conyu$conyuge == 2,
                              c("id", "edad_retro", "cor_union")],
             by = c("id", "edad_retro"),
             all.x = TRUE)
names(eder)[c(6)] <- c("cor_union2")
```

Pegado de la información de las parejas 3 a 5.

```
eder <- merge(eder, eder_conyu[eder_conyu$conyuge == 3,
                              c("id", "edad_retro", "cor_union")],
             by = c("id", "edad_retro"),
             all.x = TRUE)
names(eder)[c(7)] <- c("cor_union3")
```

```
eder <- merge(eder, eder_conyu[eder_conyu$conyuge == 4,
                              c("id", "edad_retro", "cor_union")],
             by = c("id", "edad_retro"),
             all.x = TRUE)
names(eder)[c(8)] <- c("cor_union4")
```

```
eder <- merge(eder, eder_conyu[eder_conyu$conyuge == 5,
                              c("id", "edad_retro", "cor_union")],
             by = c("id", "edad_retro"),
             all.x = TRUE)
names(eder)[c(9)] <- c("cor_union5")
```

## Fusión de eder con eder\_hijo

Con base en la forma como pegamos la información de las y los cónyuges, ahora utilizamos la función `for{}`  para pegar la información de los/as hijos/as.

```
table(eder_hijo$hijo)

##
##      1      2      3      4      5      6      7
8      9     10     11     12     13
## 51991 39527 17569 6722 3010 1367 923
556 256 206 138 70 70

for (i in 1:13) {

  conviv <- paste("conviv", i, sep = "")
  ed_hijo <- paste("ed_hijo", i, sep = "")

  eder <- merge(eder, eder_hijo[eder_hijo$hi-
jo == i,
                                c("id", "edad_
retro", "conviv", "ed_hijo")],
              by = c("id", "edad_retro"),
              all.x = TRUE)

  names(eder)[c(ncol(eder)-1, ncol(eder))] <-
c(conviv, ed_hijo)}
```

La fusión final arroja 35 variables en el objeto `eder`

```
dim(eder)

## [1] 65200 35
```

Guardamos el objeto `eder`

```
save(eder, file = "Eder.RData")
```

### 1.3.2 Construcción de la variable de interés a partir de la fusión de las bases de datos

Variable resumen de coresidencia con pareja conyugal y convivencia con los hijos

Una vez fusionadas las bases, construiremos la variable `familia` que conjuga la información de las uniones y los hijo/as de las personas entrevistadas en cada año de su vida.

```
eder$cor_union <- rowSums(eder[, c(5:9)],
na.rm = TRUE)

table(eder$cor_union, useNA = "always")

##
##      0      1 <NA>
## 40430 24770      0

eder$conviv <- rowSums(eder[, c(1:13*2+8)],
na.rm = TRUE)

table(eder$conviv, useNA = "always")

##
##      0      1      2      3      4      5      6
7      8      9     10 <NA>
## 44686 8894 7284 3034 893 252 92
18     30     15      2      0

eder$familia <- NA
eder$familia[eder$cor_union == 0 & eder$con-
viv == 0] <- 0
eder$familia[eder$cor_union == 1 & eder$con-
viv == 0] <- 1
eder$familia[eder$cor_union == 0 & eder$con-
viv != 0] <- 2
eder$familia[eder$cor_union == 1 & eder$con-
viv != 0] <- 3

table(eder$familia, useNA = "always")

##
##      0      1      2      3 <NA>
## 36979 7707 3451 17063      0
```

Etiquetamos las variables para obtener la distribución de la variable `familia` por sexo y cohorte de nacimiento.

```
eder$familia_e <- factor(eder$familia, levels
= c(0:3),
                        labels = c("No
unido/a sin hijo/as",
                                   "Unido/a
sin hijo/as",
                                   "No
unido/a con hijo/as",
                                   "Unido/a
con hijo/as"))

eder$sexo_e <- factor(eder$p2, levels =
c(1,2),
                    labels = c("Hombre",
"Mujeer"))

eder$cohorte_e <- factor(eder$cohorte, levels
= c(1:3),
                        labels = c("1948-
1952", "1968-1972", "1978-1982"))
```

Guardamos el objeto `eder` con la nueva variable generada.

```
save(eder, file = "Eder.RData")
```

### 1.3.3 Procesamiento de comandos de estadística descriptiva

Con la función `table()` obtenemos la distribución de la variable recién creada

```
table(eder$familia_e, useNA = "always")
##
## No unido/a sin hijo/as      Unido/a sin
hijo/as No unido/a con hijo/as
##                36979
7707                3451
##      Unido/a con hijo/as
<NA>
##                17063      0
```

Resultado en un momento determinado (a los 35 años)

```
table(eder$familia_e[eder$edad_retro == 35],
useNA = "always")
##
## No unido/a sin hijo/as      Unido/a sin
hijo/as No unido/a con hijo/as
##                254
175                94
##      Unido/a con hijo/as
<NA>
##                697      0
```

### Tablas de contingencia

Tabla de contingencia que combina la situación familiar con el sexo a los 35 años

```
table(eder$familia_e[eder$edad_retro == 35],
      eder$sexo_e[eder$edad_retro == 35],
useNA = "always")
##
##                Hombre Mujer <NA>
## No unido/a sin hijo/as      146  108  0
## Unido/a sin hijo/as        113   62  0
## No unido/a con hijo/as       9   85  0
## Unido/a con hijo/as         336  361  0
## <NA>                        0    0  0
```

### Resultados en porcentaje

```
tabla <- table(eder$familia_e[eder$edad_retro
== 35],
              eder$sexo_e[eder$edad_retro == 35])
```

### Porcentaje por fila

```
round(prop.table(tabla, margin = 1)*100, dig-
its = 1)
```

```
##
##                Hombre Mujer
## No unido/a sin hijo/as      57.5  42.5
## Unido/a sin hijo/as         64.6  35.4
## No unido/a con hijo/as       9.6  90.4
## Unido/a con hijo/as         48.2  51.8
```

### Porcentaje por columna

```
round(prop.table(tabla, margin = 2)*100, dig-
its = 1)
```

```
##
##                Hombre Mujer
## No unido/a sin hijo/as      24.2  17.5
## Unido/a sin hijo/as         18.7  10.1
## No unido/a con hijo/as       1.5  13.8
## Unido/a con hijo/as         55.6  58.6
```

Tabla de contingencia que combina la situación familiar con la cohorte de nacimiento a los 35 años

```
table(eder$familia_e[eder$edad_retro == 35],
      eder$cohorte_e[eder$edad_retro == 35],
useNA = "always")
##
##                1948-1952 1968-
1972 1978-1982 <NA>
## No unido/a sin hijo/as      70
84    100    0
## Unido/a sin hijo/as         45
68    62    0
## No unido/a con hijo/as      21
36    37    0
## Unido/a con hijo/as        271
218   208    0
## <NA>                        0
0      0    0
```

Incluimos el índice de origen social que se encuentra en la base `eder2019_usuarios_IOS` (`eder_ios`) y elaboramos tablas de contingencia que crucen la situación familiar con el `ios` y el sexo a los 35 años.

```

eder_ios <- read.table("eder2019_usuarios_
IOS.txt",
                    header = TRUE, sep = ";")

eder<- merge(eder, eder_ios[, c("id", "ios-
terciles")],
            by = "id", all.x = TRUE)

eder$ios_e <- factor(eder$ios_terciles, levels
= c(1:3),
                  labels = c("Bajo", "Medio", "Alto"))

table(eder$familia_e[eder$edad_retro == 35],
      eder$ios_e[eder$edad_retro == 35],
      eder$sexo_e[eder$edad_retro == 35])

## , , = Hombre
##
##
##           Bajo Medio Alto
## No unido/a sin hijo/as   40  57  48
## Unido/a sin hijo/as     34  39  40
## No unido/a con hijo/as   3   3   3
## Unido/a con hijo/as    122 115  95
##
## , , = Mujer
##
##
##           Bajo Medio Alto
## No unido/a sin hijo/as   27  43  37
## Unido/a sin hijo/as     15  21  25
## No unido/a con hijo/as   43  24  17
## Unido/a con hijo/as    145 107 106

```

## Gráficas

Con las tablas de contingencia pudimos observar la distribución de la variable sobre la situación familiar para una edad determinada (35 años). Con las gráficas, podemos reconstruir la historia de los eventos para un período temporal (en este ejemplo, entre los 20 y los 40 años). Adicionalmente, lo compararemos entre cohortes de nacimiento para poder analizar los cambios en el tiempo histórico.

Para generar un gráfico de la distribución de la situación familiar por cohorte de nacimiento y edad usamos la librería ggplot2.

```

install.packages("ggplot2")

library(ggplot2)

## Warning: package 'ggplot2' was built under
R version 4.3.1

```

Primero debemos preparar los datos. Generamos una variable auxiliar para el conteo de casos.

```
eder$aux <- 1
```

Creamos el objeto datos que agrega la suma de casos por cada categoría de la variable sobre la situación familiar para cada cohorte y cada año retrospectivo.

```

datos <- aggregate(list(eder$aux),
                  by=list(eder$cohorte,
                          eder$edad_retro, eder$familia),
                  FUN=sum,
                  na.rm=TRUE)
names(datos) <- c("Cohorte", "Edad", "Familia",
"Suma")

```

Creamos el objeto datos2 que agrega la cantidad de casos para cada cohorte y año retrospectivo.

```

datos2 <- aggregate(list(eder$aux),
                   by = list(eder$cohorte,
                              eder$edad_retro),
                   FUN = sum,
                   na.rm = TRUE)

names(datos2) <- c("Cohorte", "Edad", "Total")

```

Fusionamos los objetos datos y datos2.

```
datos <- merge(datos, datos2, by = c("Cohorte", "Edad"), all = TRUE)
```

Calculamos el porcentaje.

```
datos$Porcentaje <- datos$Suma/datos$Total*100
```

Etiquetamos las variables Familia y Cohorte.

```

datos$Familia <- factor(datos$Familia, levels
= c(0:3),
                      labels = c("No unido sin hijo/as",
                                  "Unido sin hijo/as",
                                  "No unido con hijo/as",
                                  "Unido con hijo/as"))

datos$Cohorte <- factor(datos$Cohorte, levels
= c(1:3),
                        labels = c("1948-
1952", "1968-1972", "1978-1982"))

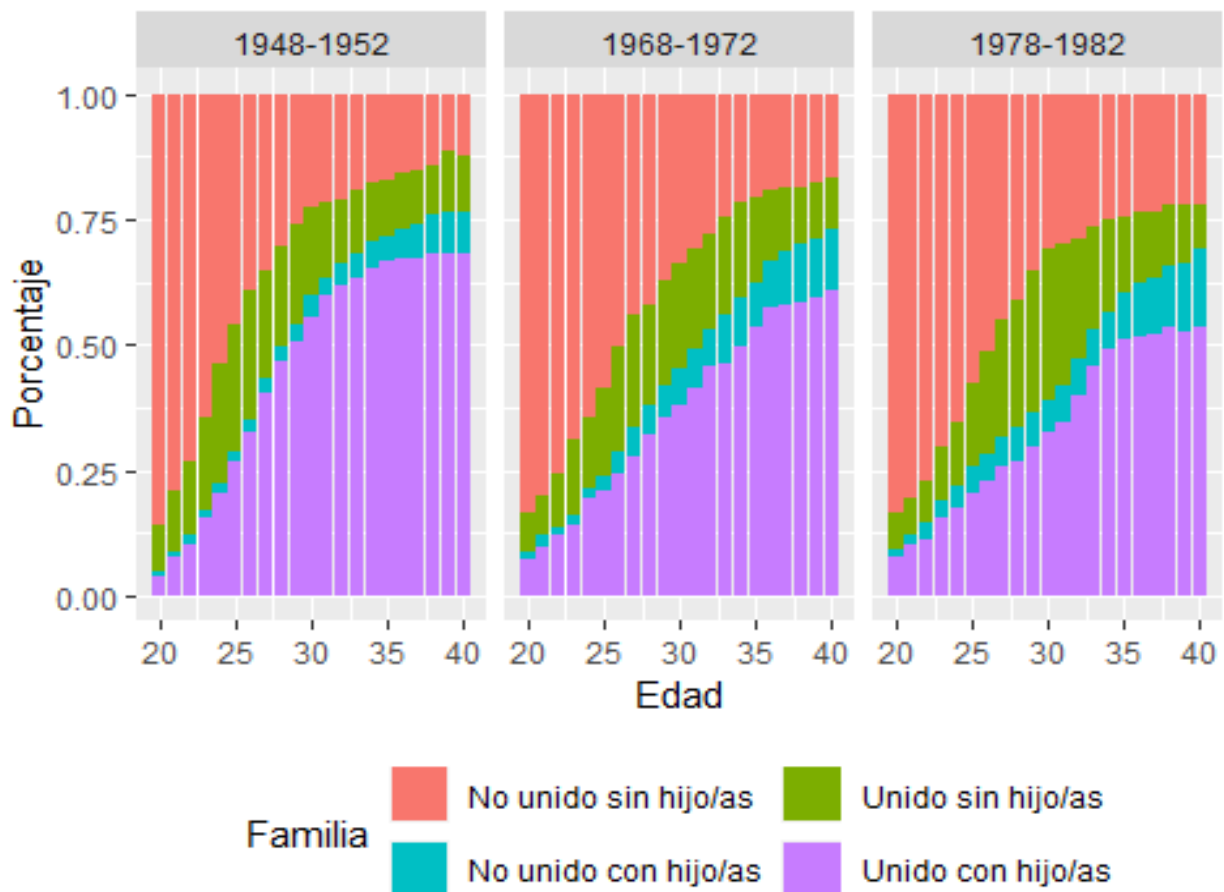
```

Ahora generamos el gráfico de barras apiladas.

```

ggplot(datos[datos$Edad %in% c(20:40),],
      aes(fill = Familia, y = Porcentaje, x =
Edad)) +
  geom_bar(position = "fill", stat = "identi-
ty") +
  facet_wrap(~Cohorte) +
  theme(legend.position = "bottom") +
  guides(fill = guide_legend(title="Familia",
nrow = 2, byrow = TRUE))

```



En el gráfico, cada barra representa la distribución de las categorías de la variable `Familia` entre los 20 y los 40 años para las tres cohortes de nacimiento. Puede observarse el contraste entre las cohortes de los extremos que da cuenta de los cambios en la conformación de la familia a lo largo del tiempo. Mientras que la cohorte más envejecida (1948-1952) predomina el estado “Unido con hijo/as”, en la cohorte más joven (1978-1982) cobra más relevancia el estado “No unido sin hijo/as”.

Para guardar el gráfico en la carpeta de trabajo usamos la función `ggsave()`.

```
ggsave("familia_cohorte.png", plot = last_
plot(), device = "png", dpi = 300)

## Saving 5 x 4 in image
```

## 2. ¿Qué técnicas estadísticas longitudinales se pueden aplicar con la EDER-CABA?

### 2.1 Aproximación teórica y ejemplos de aplicación

#### 2.1.1 Tabla de vida (con casos truncados)

La Tabla de Vida es un modelo que mide las probabilidades de ocurrencia y supervivencia de los eventos de una población en función de la edad.

Caracteriza los eventos identificando las edades más frecuentes en las que tienen lugar y qué otros atributos (por ejemplo, el género, el origen social o la etnia, entre otros) colocan a la población en una situación de relativa “ventaja” o “desventaja” respecto de experimentar dicho evento.

#### Glosario

**Evento:** Acontecimiento que ocurre en un momento determinado, tal como el matrimonio, nacimiento de un hijo, cambio de lugar de residencia.

Los eventos registrados en la EDER-CABA son los de duración de al menos 12 meses (un año calendario).

**Probabilidad de ocurrencia:** Es el estimador de la probabilidad condicional de ocurrencia del evento durante cualquier intervalo dada la exposición al riesgo de ocurrencia al inicio de dicho intervalo (ej., la probabilidad de unirse entre las edades exactas ‘x’ ‘y’).

**Probabilidad de supervivencia:** Es la proporción condicional que sobrevive al evento (ej., entrar en unión) durante el intervalo, dada la exposición al riesgo de ocurrencia al inicio de dicho intervalo.

El principal potencial de esta técnica consiste en incorporar la noción dinámica del tiempo ya que identifica cuánto tiempo ha tardado cada persona en alcanzar los eventos bajo estudio (Martínez y De Irala, 2005).

A partir de esta información pueden aplicarse otras métricas como las proporciones acumuladas por edades de ocurrencia de un evento, el tiempo promedio esperado de permanencia en un estado, cuartiles, entre otras medidas resúmenes de intensidad y calendario.

Para consultar ejemplos sobre proporciones acumuladas en base a las Tablas de Vida, consulte el [Informe metodológico y primeros resultados](#) disponible en el sitio web de la DGEyC.

En suma, los principales alcances de las Tablas de Vida (Gómez Melis, 2002) consisten en:

- La descripción y resumen de los tiempos hasta el evento, a partir de la estimación de las funciones de supervivencia y probabilidades de ocurrencia por edades.
- La comparación de las distribuciones de los tiempos de supervivencia correspondientes a una misma población diferenciada de acuerdo a algún atributo (variable explicativa).
- El establecimiento y comprensión de la relación que pueda existir entre los tiempos hasta el evento y esta variable explicativa.

Entre algunas de las aplicaciones de la Tabla de Vida en base a los datos de la EDER-CABA, se destaca que se puede analizar el calendario de ocurrencia de eventos tales como la ma/paternidad, la salida del hogar de origen y la entrada en unión, entre otros.

En estos casos, la “supervivencia” refiere a las personas que hasta una determinada edad no han experimentado dichos eventos.

Las tablas que se utilizan con este tipo de información retrospectiva como los de la EDER-CABA son las Tablas de Vida con casos truncados.

#### Glosario

**Casos truncados:** Un caso es “truncado” cuando no experimenta el evento durante el tiempo de seguimiento y, por lo tanto, se desconoce cuál es su tiempo de supervivencia.

El truncamiento de la información es una situación común cuando se dispone de datos reales de tiempo al evento (Solís, 2011). Así, las Tablas de Vida con casos truncados permiten incorporar el tiempo de todas las personas de la muestra (quienes experimentaron el evento y quienes no lo han experimentado)

Con la EDER-CABA trabajamos con “truncamiento por la derecha”, es decir, personas que al momento de la encuesta pertenecen al conjunto en riesgo, pero que no han experimentado el evento en cuestión.

Las Tablas de Vida tienen algunas limitaciones que llevan a la necesidad de utilizar técnicas estadísticas más avanzadas (como las regresiones logísticas de tiempo discreto que veremos a continuación):

- Las Tablas de Vida son “consumidoras de casos”, es decir, requieren subdivisiones progresivas de los datos para establecer controles simultáneos por cada variable independiente. Esto genera que, al incorporar nuevas variables, pronto se llegue a una situación en que no se cuentan con datos suficientes para realizar el análisis.
- Resulta complejo evaluar si el efecto de una variable independiente es estadísticamente significativo una vez que se controla por los efectos de otras variables (Solís, 2011).

### Ejemplo de aplicación

El siguiente ejemplo<sup>2</sup> muestra una tabla resumen con dos estadísticos descriptivos de análisis de supervivencia resultantes de la aplicación de la Tabla de Vida con casos truncados: cuantiles y proporciones acumuladas de ocurrencia del evento por edad según sexo (1-Sedad).

Los resultados derivan de dos encuestas transversales de las cuales se tomó la información retrospectiva (preguntas sobre la temporalidad de los eventos, en

este ejemplo, la edad en la que dejaron la escuela): Encuesta Anual de Hogares de la Ciudad de Buenos Aires de 2010 (EAH) y Encuesta Nacional de la Juventud de 2010 para el Distrito Federal (ENJ).

El cuadro muestra el calendario de salida de la escuela de los jóvenes en la Ciudad Autónoma de Buenos Aires (CABA) y en el Distrito Federal, de acuerdo con el sexo y el grupo de edad (20 a 24 y 25 a 29 años).

Con base en los cuantiles, se observa que en la capital mexicana los jóvenes de los dos grupos de edad salen de la escuela a edades más tempranas que los de la capital argentina. Las mujeres porteñas tienen un calendario más temprano de salida de la escuela, a diferencia de las porteñas, que permanecen más tiempo en el sistema escolar.

Las medidas parciales de intensidad confirman lo antes expuesto ya que se observa un calendario más temprano en la salida de la escuela en el Distrito Federal. Se observa una diferencia marcada entre los sexos, con un comportamiento inverso entre jurisdicciones.

En CABA las mujeres dejan la escuela más tarde que los varones, mientras que entre los jóvenes porteños ocurre lo contrario.

Entre los porteños del grupo de edad 20-24 casi 38% salieron de la escuela a los 20 años, frente a un 29% de las mujeres. En cambio, en la capital mexicana el 51% de las mujeres y el 34% de los varones ya habían salido de la escuela a esa edad.

Con relación a la edad, hacia los 20 años casi no se perciben diferencias en el calendario de salida de la escuela en las dos ciudades.

No obstante, entre los varones de 25 a 29 años se observa una proporción mayor de varones en el Distrito Federal que dejan la escuela más tempranamente que sus pares porteños.

2 Puede ampliar el desarrollo de este ejemplo en el artículo [“Entre la escuela y el trabajo. El tránsito a la vida adulta de los jóvenes en la Ciudad de Buenos Aires y el Distrito Federal”](#) (Autores: Sabrina Ferraris y Mario Martínez Salgado, 2015).

Fig. nº2. Ejemplo de aplicación Tabla de Vida con casos truncados

Calendario de la salida de la escuela de los jóvenes de la Ciudad de Buenos Aires y del Distrito Federal, 2010									
Cuantil	Buenos Aires				Distrito Federal				
	20 a 24 años		25 a 29 años		20 a 24 años		25 a 29 años		
	H	M	H	M	H	M	H	M	
10	17.01	17.65	16.08	17.51	15.9	15.2	16.4	14.4	
25	18.72	20.12	18.48	19.52	18.5	17.0	18.1	16.2	
50	23.42*	24.16*	23.90	25.25	22.3*	19.8*	21.7	18.5	
Medidas parciales de intensidad									
1-S <sub>20</sub>	0.3793	0.2845	0.3965	0.2890	0.34050	0.51670	0.3951	0.6483	
1-S <sub>23</sub>	0.5246*	0.4789*	0.5030	0.4240	0.59080*	0.69630*	0.5888	0.7890	
1-S <sub>25</sub>	—	—	0.6227	0.5710	—	—	0.7514	0.8671	
1-S <sub>28</sub>	—	—	0.7697	0.7773	—	—	0.8688	0.9339	

\* Las estimaciones del grupo 20 a 24 años podrían estar afectadas por el alto número de casos truncados.

### 2.1.2 Modelos de Regresión de Tiempo Discreto

Una alternativa para superar las limitaciones de las Tablas de Vida son los Modelos de Regresión de Tiempo Discreto.

Estos modelos permiten ver el efecto de distintas variables independientes en la probabilidad de experimentar el evento, controlado simultáneamente por el patrón de dependencia temporal y otras variables que se consideren pertinentes (Solís, 2011).

Los principales alcances de los Modelos de Regresión de Tiempo Discreto son:

- Predice la probabilidad de ocurrencia o no del evento, como se suele realizar con la técnica de regresión logística.
- También permite predecir el cuándo, es decir, su distribución de acuerdo al calendario.
- Incorpora variables independientes tanto fijas como móviles o cambiantes en el tiempo.

Esto habilita una perspectiva dinámica y de procesos que no es accesible mediante el uso de técnicas convencionales basadas en una aproximación sincrónica (Solís, 2011).

#### Glosario

**Variables fijas en el tiempo:** Son aquellas características de las personas que se encuentran determinadas en el momento en que comenzamos a observarlas y no se modifican hasta que salen del conjunto en riesgo (ej., la cohorte de nacimiento).

**Variables móviles o cambiantes en el tiempo:** Son aquellas que refieren a características que pueden modificarse durante el tiempo de exposición al riesgo (ej., cambios en las uniones conyugales).

Resulta conveniente utilizar esta técnica cuando se cuenta con datos en edades “cumplidas”, es decir, sobre la ocurrencia de los eventos en intervalos de duraciones.

En este caso, el riesgo instantáneo (‘hazard’) no puede ser calculado directamente, y debe ser aproximado mediante la estimación de la probabilidad de experimentar el evento para el intervalo de tiempo en cuestión (Allison, 1984; Solís, 2011).

## Ejemplo de aplicación

A continuación, se presenta un ejemplo de aplicación de los Modelos de Regresión de Tiempo Discreto elaborados a partir del evento de primera salida de la escolarización formal como variable dependiente para ambos sexos, en base a los datos de la EDER-CABA 2019<sup>3</sup>.

Fundamentalmente, estos modelos permiten observar las chances de ocurrencia (*odds ratio* o razones de momios), y el cuándo, de un evento en particular en función de un conjunto de variables, que entre ellas pueden estar, por ejemplo, los calendarios de varios eventos (las concurrencias entre eventos).

En este ejemplo, se analizan las chances de vivenciar el evento identificado como variable dependiente (primera salida de la escolarización), en varones y mujeres según las variables independientes (la cohorte de nacimiento, el clima educativo del hogar de origen, el primer empleo y el ingreso a la ma/paternidad).

Se presentan los resultados bajo una lógica de “modelos anidados”. En el primer modelo se incluye una sola variable independiente y en los modelos subsiguientes se van agregando nuevas variables, con el fin de identificar si los efectos de las primeras variables se mantienen al controlar por otros factores y, del mismo modo, observar el efecto de estos nuevos factores sobre el evento bajo análisis.

Los datos destacados en la tabla indican la interdependencia de las esferas de la actividad educativa, laboral y familiar. En efecto, si tanto varones como mujeres comienzan a trabajar, luego de los dos primeros años de haber transitado el evento, tienen las mismas chances de salir de la escolarización. Además, si tienen un hijo/a, las mujeres tienen chances mucho más marcadas –casi el doble– que los varones de salir de la escolarización (5,12 vs. 2,69).

3 Puede ampliar el desarrollo de este ejemplo en el artículo “[Desigualdades de género en las transiciones a la adultez en la Ciudad Autónoma de Buenos Aires, EDER-2019](#)” (Autora: Mariela Giacomponello, 2023).

Fig. n°3. Ejemplo de aplicación Modelo de Regresión de Tiempo Discreto

**Factores asociados a la primera salida de la escolarización formal. Razones de momios de modelos logísticos de tiempo discreto según género, EDER-CABA (2019)**

	Varones				Mujeres			
	M1	M2	M3	M4	M1	M2	M3	M4
<b>Tramos de edad (Ref. = 10-12)</b>								
13-14	1,39	1,46*	1,44*	1,44*	1,23	1,27	1,26	1,25
15-16	2,70***	3,00***	2,89***	2,89***	3,42***	3,63***	3,58***	3,56***
17-18	7,99***	9,48***	8,79***	8,78***	9,15***	10,32***	9,73***	9,78***
19-20	3,84***	4,88***	4,21***	4,20***	6,73***	7,99***	6,93***	6,95***
21-23	6,39***	8,79***	6,95***	6,92***	11,67***	14,65***	11,64***	11,62***
24-25	13,63***	19,06***	14,51***	14,54***	21,60***	27,65***	20,93***	20,71***
26-28	9,04***	11,72***	8,56***	8,15***	21,62***	28,76***	21,50***	20,09***
29-30	7,34***	8,76***	6,39***	5,96***	43,38***	59,12***	43,43***	37,71***
<b>Generación (Ref. = 1948-1952)</b>								
1968-1972	0,93	1,17	1,16	1,17	0,70***	0,84	0,82	0,82
1978-1982	0,83	1,1	1,11	1,11	0,63***	0,78**	0,78**	0,77**
<b>Clima educativo del hogar de origen (Ref. = bajo)</b>								
CEHO-medio		0,49***	0,50***	0,50***		0,67**	0,66**	0,68**
CEHO-alto		0,24***	0,25***	0,25***		0,32***	0,32***	0,33***
<b>Transición laboral (Ref. = No trabaja)</b>								
Primer trabajo transicional (hasta 2 años)			1,13	1,13			1,21	1,22
Primer trabajo postransicional (después de 2 años)			1,40**	1,40**			1,42**	1,41**
<b>Transición mater/paternidad (Ref. = Sin hijo/as)</b>								
Primer hijo/a transicional (hasta 2 años)				0,81				0,77
Primer hijo/a postransicional (después de 2 años)				2,69*				5,12***
Pseudo R2	0,1	0,12	0,12	0,12	0,13	0,15	0,15	0,15
Años-persona	5761	5761	5761	5761	5473	5473	5473	5473

Significación de las variables: \* p<.1; \*\* p<.05; \*\*\* p<.01

### 2.1.3 Análisis de Secuencias

El Análisis de Secuencias busca identificar patrones de información en los calendarios de ocurrencia de los eventos (cambios de estado), dando lugar a conocer qué tan diversos o semejantes son esos itinerarios y, a partir de esto, delimitar grupos de la población o tipologías (Gauthier et al., 2014).

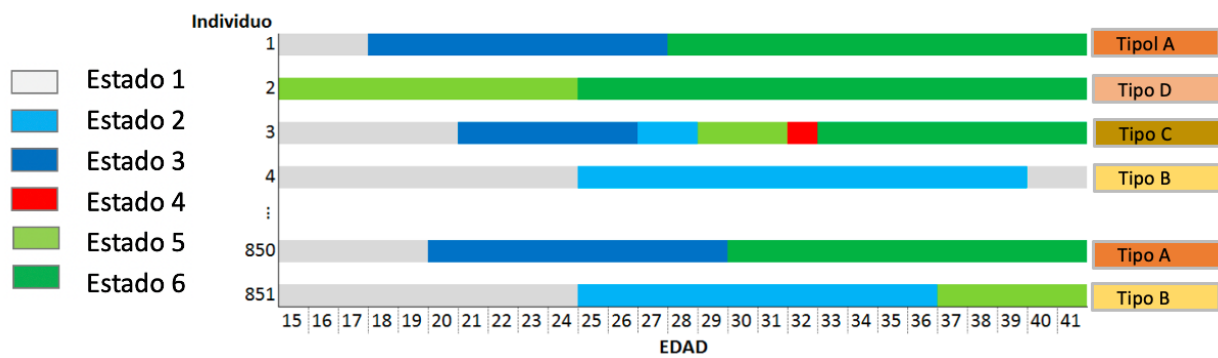
#### Glosario

**Secuencia:** Es una lista ordenada de estados. Los elementos que la componen (número de estados, cambios de estados y patrones de frecuencias ) y su longitud son funciones del tiempo (Brzinsky-Fay y Kohler, 2010).

El foco está puesto en el orden de los estados definidos de antemano y en su sucesión cronológica, con la intención de evidenciar que la realidad social ocurre en historias más que en instantes (Brzinsky-Fay y Kohler, 2010).

En el ejemplo en la tabla a continuación, puede verse cómo cada barra representa un caso (individuos). Los estados se distribuyen a lo largo del calendario por edades y pueden agruparse en tipologías por similitud (Tipo A, B, C y D).

Fig. nº4. Visualización de secuencia de estados e identificación de grupos por similitud



Entre las técnicas que forman parte del Análisis de Secuencias, el análisis de alineación óptima (OMA por sus siglas en inglés: Optimal Matching Analysis) permite encontrar patrones en las secuencias con base en una medida de la proximidad o semejanza entre ellas.

El principal procedimiento del OMA consiste en comparar dos secuencias de estados con el fin de identificar que tan semejantes son. Esta tarea se realiza alineando las secuencias por pares y se transforma una secuencia en la otra a partir de inserciones, borrados y sustituciones de estados.

Al utilizar cada una de las operaciones se incurre en un costo, por lo que el objetivo en esta parte del proceso es obtener la transformación entre pares de secuencias con la suma de costos mínima (para transformar una secuencia en otra es probable que se tenga que utilizar más de una operación).

El resultado de este procedimiento se conoce como matriz de distancias y sirve de insumo para que con un análisis de cluster se construya una tipología de grupos de secuencias (Abbot y Tsay, 2001; Gauthier et. al., 2014).

La agrupación es la parte técnica del análisis, que implica la conformación de una tipología (de varios tipos) hecha de características individuales que son tan similares como sea posible dentro de un grupo y tan disímiles como sea posible entre grupos (Gauthier et. al., 2014).

En suma, el análisis de secuencias consiste centralmente en cuatro etapas:

- Construcción de secuencias individuales de estados.
- Comparación de secuencias según su grado de disimilitud.
- Agrupamiento de las secuencias más similares para reducir la complejidad y crear una tipología.
- Contextualización de la tipología obtenida con variables externas de interés (ej., sexo, generaciones, grupos sociales, etc.) (Gauthier et. al., 2014).

## Ejemplo de aplicación

En este ejemplo se presentan un conjunto de gráficas como resultado de la aplicación del Análisis de Secuencias en base a los la EDER-México 2011.<sup>4</sup>

Se analiza la forma en que los hombres mexicanos construyen sus trayectorias de proveeduría a lo largo del curso de vida de dos generaciones que han vivido contextos socioeconómicos diversos (1951-53 y 1966-68).

A partir del tipo de empleo (formal o informal) y los periodos donde se es el principal sostén económico del hogar, se construyeron las secuencias e identificaron seis trayectorias de proveeduría.

En la gráfica puede observarse que en los primeros cuatro se destacan lapsos prolongados de proveeduría económica.

La trayectoria 1 (T1) representa una proveeduría temprana con empleo en la economía formal en buena parte de su trayectoria.

En el segundo grupo (T2), se observa una proveeduría tardía con empleo en la economía formal.

La trayectoria 3 (T3) se caracteriza por una proveeduría con alternancia en el tipo de empleo (formal/informal), con un calendario del inicio temprano (alrededor de los veinte años).

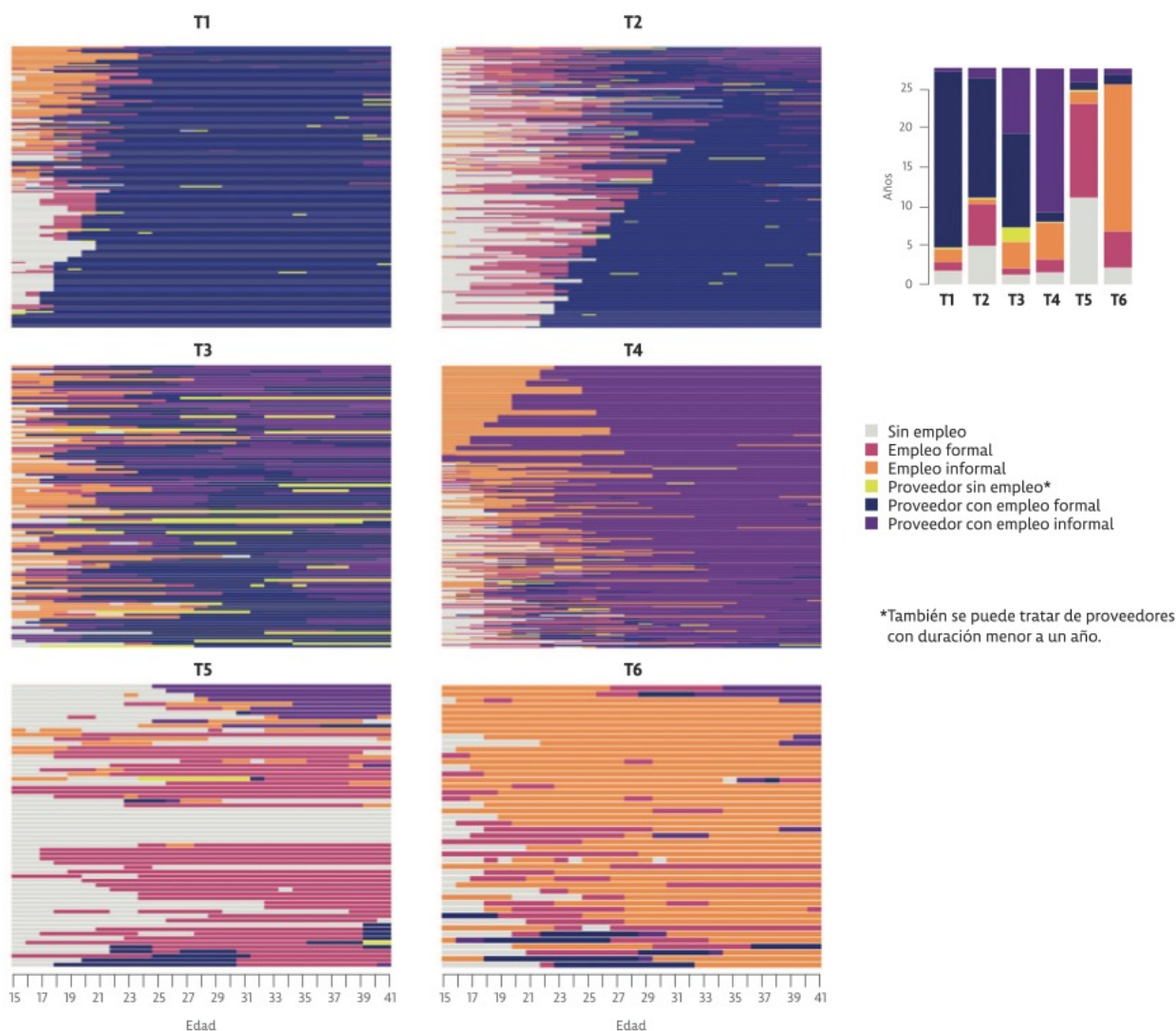
En el cuarto grupo (T4) predomina una proveeduría con empleo en la economía informal.

Los dos tipos restantes de trayectorias se caracterizan por periodos sin empleo (o en empleos de corta duración), seguido de tramos con trabajo pero sin proveeduría.

<sup>4</sup> Puede ampliar el desarrollo de este ejemplo en el artículo [“Ahí te dejo esos dos pesos”. Trayectorias de proveeduría económica de los hombres en México”](#) (Autores: Mario Martínez Salgado y Sabrina Ferraris, 2023).

Fig. nº5. Ejemplo de aplicación Análisis de Secuencias

### Trayectorias de proveeduría económica y tiempo promedio en cada estado. México, 2011



### 3. Bibliografía citada

**Abbot, A. & Tsay, A.** (2001). Sequence Analysis and Optimal Matching Methods in Sociology. Review and Prospect. *Sociological Methods and Research*, 29(1): 3-33.

**Allison, P. D.** (1984). *Event history analysis. Regression for Longitudinal Event Data*. Beverly Hills, Calif: Sage Publications.

**Brzinsky-Fay, C. & U. Kohler** (2010). New Developments Sequence Analysis. *Sociological Methods & Research*. Sage, vol. 38, núm. 3, pp. 359-512.

**DGEyC** (2021). EDER CABA 2019. Informe metodológico y primeros resultados. *Dirección General de Estadística y Censos*.

**Ferraris, S. A., & Martínez Salgado, M.** (2015). Entre la escuela y el trabajo. El tránsito a la vida adulta de los jóvenes en la Ciudad de Buenos Aires y el Distrito Federal. *Estudios Demográficos Y Urbanos*, 30(2), 405–431. <https://doi.org/10.24201/edu.v30i2.1479>

**Gauthier, J., Bühlmann, F. & Blanchard, P.** (2014). “Introduction: Sequence Analysis in 2014”. En Blanchard, Bühlmann y Gauthier (eds.), *Advances in Sequence Analysis: Theory, Method, Applications*, Springer, pp. 1-17.

**Giacoponello, M.** (2023). Desigualdades de género en las transiciones a la adultez en CABA, EDER-2019. *Inflexiones. Revista de Ciencias Sociales y Humanidades*. México: Editorial UNAM. <http://dx.doi.org/10.22201/udir.2954341xp.164>

**Gómez Melis, G.** (2002). Técnicas estadísticas en Análisis de Supervivencia. *ICB digital* N°5. Recuperado de: [www.icbdigital.org/icbdigital/](http://www.icbdigital.org/icbdigital/)

**Martínez González, M. A. & De Irala Estévez, J.** (2005). “Análisis de supervivencia y análisis multivariado”. En F. López-Jiménez. G.T. Obrador Vera. México, 2a Edición: Manual Moderno.

**Salgado, M. & Ferraris, S.** (2023). “‘Ahi te dejo esos dos pesos’. Trayectorias de proveeduría económica de los hombres en México”. En M. Mier y Terán (coord.) *Trayectorias y desigualdades sociales en el contexto mexicano: una perspectiva longitudinal*. México: UNAM, Instituto de Investigaciones Sociales. <https://ru.iis.sociales.unam.mx/handle/IIS/6086>

**Solís, P.** (2011). “Guía práctica de análisis de historia de eventos en Stata”, Centro de Estudios Sociológicos El Colegio de México (mimeo).

### Bibliografía ampliatoria

**Coubès, M. L., Solís, P. & Zavala de Cosío, M. E.** (coords.) (2016). *Generaciones, cursos de vida y desigualdad social en México*. El Colegio de México y El Colegio de la Frontera norte, México.

**Coubès, M. L., Zavala de Cosío, M. E. & Zenteno, R.** (coords.) (2005). *Cambio demográfico y social en el México del siglo XX. Una perspectiva de historia de vida*. México: El Colegio de la Frontera Norte.

**Courgeau, D. & Lelièvre, E.** (2001). *Análisis demográfico de las biografías*. México D.F.: El Colegio de México.

**Mier y Terán, M.** (coord.) (2023). *Trayectorias y desigualdades sociales en el contexto mexicano: una perspectiva longitudinal*. México: Editorial UNAM, Instituto de Investigaciones Sociales. <https://ru.iis.sociales.unam.mx/handle/IIS/6086>

**Zavala, M. E.** (coord.) (2023). *La Odisea de las generaciones: de las historias de vida a los territorios*. El Colegio de México.

**Zavala, M. E., Lago, M. E., Olmos, M. F., & Aguilera, M. E.** (2021). Transition to adulthood in Mexico and Buenos Aires cities: A retrospective demographic approach of three generations. *Revista Latinoamericana De Población*, 16, e202121. <https://doi.org/10.31406/relap2022.v16.e202121>

Para más información estadística podés seguirnos en nuestras redes,  
visitarnos en nuestra web o consultarnos por mail.