Data preparation

Changes in the Justification of Pension Inequality in Chile (2016–2023) and its Relationship to Social Class and Beliefs in Meritocracy

Andreas Laffert, Research Assistant

2025-09-11

1 Presentation

This is the data preparation code for the paper "Changes in the Justification of Pension Inequality in Chile (2016–2023) and its Relationship to Social Class and Beliefs in Meritocracy". The prepared dataset is `ELSOC_Long_2016_2023_1.00.RData`.

2 Libraries

```
if (! require("pacman")) install.packages("pacman")

pacman::p_load(tidyverse,
               car,
               sjmisc,
               here,
               sjlabelled,
               SciViews,
               naniar,
               readxl,
               sjPlot)
```

```r
options(scipen=999)
rm(list = ls())
```

3 Data

```r
load(url("https://dataverse.harvard.edu/api/access/datafile/10797987"))

glimpse(elsoc_long_2016_2023)
```

4 Processing

```r
elsoc_long_2016_2023[elsoc_long_2016_2023 ==-999] <- NA
elsoc_long_2016_2023[elsoc_long_2016_2023 ==-888] <- NA
elsoc_long_2016_2023[elsoc_long_2016_2023 ==-777] <- NA
elsoc_long_2016_2023[elsoc_long_2016_2023 ==-666] <- NA

elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(just_pension = d02_01,
         merit_effort = c18_09,
         merit_talent = c18_10,
         age = m0_edad, m01,
         sex = m0_sexo,
         ideo = c15) %>%
  as_tibble() %>%
  sjlabelled::drop_labels(., drop.na = FALSE)
```

```r
# Market Justice Preferences

frq(elsoc_long_2016_2023$just_pension)
```

```
Grado de acuerdo: Justicia distributiva en pensiones (x) <numeric>
# total N=20761 valid N=17966 mean=2.24 sd=1.11
```

Value	Label	N	Raw %	Valid %	Cum. %
1	Totalmente en desacuerdo	4889	23.55	27.21	27.21
2	En desacuerdo	7802	37.58	43.43	70.64
3	Ni de acuerdo ni en desacuerdo	1740	8.38	9.68	80.32
4	De acuerdo	3087	14.87	17.18	97.51
5	Totalmente de acuerdo	448	2.16	2.49	100.00
<NA>	<NA>	2795	13.46	<NA>	<NA>

```
elsoc_long_2016_2023$just_pension <- car::recode(elsoc_long_2016_2023$just_pension,
recodes = c("1='Strongly disagree'; 2='Disagree'; 3='Neither agree nor disagree'; 4='A
levels = c("Strongly disagree", "Disagree", "Neither agree nor disagree", "Agree", "St
as.factor = T)

elsoc_long_2016_2023$just_pension <- sjlabelled::set_label(elsoc_long_2016_2023$just_p
                         label = "Pension distributive justice")
```

```
# Social class scheme EOW

frq(elsoc_long_2016_2023$m07)
```

```
Relacion de empleo (x) <numeric>
# total N=20761 valid N=7288 mean=2.35 sd=1.90

Value
-----
    1
    2
    3
    4
    5
    6
    7
 <NA>

                                                                                Label
---------------------------------------------------------------------------------------
                                                Empleado u obrero en empresa privada
     Empleado u obrero del sector publico (incluso empresa publica o municipalidad)
                                                Miembro de las Fuerzas Armadas y de Orden
```

Patron/a o empleador/a (contrata o paga a honorarios a uno/o o mas trabajadores/as)
Trabaja solo, no tiene empleados/as
Familiar no remunerado
Servicio domestico
<NA>

N	Raw %	Valid %	Cum. %
4160	20.04	57.08	57.08
964	4.64	13.23	70.31
80	0.39	1.10	71.41
329	1.58	4.51	75.92
1379	6.64	18.92	94.84
15	0.07	0.21	95.05
361	1.74	4.95	100.00
13473	64.90	<NA>	<NA>

```
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(rel_empleo = factor(m07,
                             levels = 1:7,
                             labels= c("Empleado u obrero en empresa privada",
                                       "Empleado u obrero del sector público",
                                       "Miembro de las Fuerzas Armadas y de Orden",
                                       "Patrón/a o empleador/a",
                                       "Trabaja solo, no tiene empleados",
                                       "Familiar no remunerado",
                                       "Servicio doméstico"
                                       )))


# Definir los niveles y etiquetas originales
niveles_rel_empleo <- 1:7
etiquetas_rel_empleo <- c("Empleado u obrero en empresa privada",
                          "Empleado u obrero del sector público",
                          "Miembro de las Fuerzas Armadas y de Orden",
                          "Patrón/a o empleador/a",
                          "Trabaja solo, no tiene empleados",
                          "Familiar no remunerado",
                          "Servicio doméstico")
```

```r
# Crear columna desplazada y rellenar valores NA
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  group_by(idencuesta) %>%                                          # Agrupa por i
  mutate(rel_empleo_lagged = lag(as.character(rel_empleo), n = 1)) %>%  # Desplaza rel
  ungroup() %>%
  mutate(rel_empleo = if_else(!is.na(rel_empleo),
                              as.character(rel_empleo),
                              rel_empleo_lagged),
         rel_empleo = factor(rel_empleo,                            # Convertir de
                             levels = etiquetas_rel_empleo)) %>%
  select(-rel_empleo_lagged)                                        # Elimina la c

# Tabla de frecuencias y porcentajes
sjt.xtab(elsoc_long_2016_2023$rel_empleo,elsoc_long_2016_2023$ola,
         show.col.prc=TRUE,
         var.labels=c("Relación de empleo","Ola"),
         show.summary=FALSE,         title="Frecuencias y porcentajes de Relación de E
```

Table 1: Frecuencias y porcentajes de Relación de Empleo, por ola

Relación de Empleo	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	
Empleado u obrero en empresa privada	1091 61.1 %	898 60.3 %	1349 58.3 %	1147 57.8 %	897 54.1 %	731 52.8 %	866 54 %	6979 57.1 %
Empleado u obrero del sector público	186 10.4 %	155 10.4 %	291 12.6 %	262 13.2 %	229 13.8 %	195 14.1 %	264 16.5 %	1582 12.9 %

Miembro de las Fuerzas Armadas y de Orden	25 1.4 %	17 1.1 %	25 1.1 %	21 1.1 %	19 1.1 %	11 0.8 %	12 0.7 %	130 1.1 %
Patrón/a o empleador/a	85 4.8 %	72 4.8 %	109 4.7 %	92 4.6 %	86 5.2 %	67 4.8 %	55 3.4 %	566 4.6 %
Trabaja solo, no tiene empleados	321 18 %	277 18.6 %	415 17.9 %	354 17.8 %	347 20.9 %	307 22.2 %	309 19.3 %	2330 19.1 %
Familiar no remunerado	4 0.2 %	3 0.2 %	6 0.3 %	4 0.2 %	1 0.1 %	2 0.1 %	4 0.2 %	24 0.2 %
Servicio doméstico	75 4.2 %	67 4.5 %	118 5.1 %	106 5.3 %	78 4.7 %	71 5.1 %	94 5.9 %	609 5 %
Total	1787 100 %	1489 100 %	2313 100 %	1986 100 %	1657 100 %	1384 100 %	1604 100 %	12220 100 %

```
elsoc_long_2016_2023$rel_empleo2 <- car::recode(elsoc_long_2016_2023$rel_empleo,
                                  recodes=c("'Patrón/a o empleador/a'='Empleador';
                                            'Trabaja solo, no tiene empleados'='Autoemp
                                            'Empleado u obrero en empresa privada'='Asa
                                            'Empleado u obrero del sector público'='Asa
                                            'Miembro de las Fuerzas Armadas y de Orden'
                                            'Familiar no remunerado'='Asalariado';
                                            'Servicio doméstico'='Asalariado'"),
                                  as.factor=TRUE, # convertir a factor
                                  levels=c("Empleador",
                                           "Autoempleado",
                                           "Asalariado")) # ordenar niveles

# Tabla de frecuencias y porcentajes
sjt.xtab(elsoc_long_2016_2023$rel_empleo2,elsoc_long_2016_2023$ola,
         show.col.prc=TRUE,
```

```
          var.labels=c("Relación de empleo 2","Ola"),
          show.summary=FALSE,
          title="Frecuencias y porcentajes de Relación de Empleo 2, por ola")
```

Table 2: Frecuencias y porcentajes de Relación de Empleo 2, por ola

Relación de Empleo 2	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	
Empleador	85	72	109	92	86	67	55	566
	4.8 %	4.8 %	4.7 %	4.6 %	5.2 %	4.8 %	3.4 %	4.6 %
Autoempleado	321	277	415	354	347	307	309	2330
	18 %	18.6 %	17.9 %	17.8 %	20.9 %	22.2 %	19.3 %	19.1 %
Asalariado	1381	1140	1789	1540	1224	1010	1240	9324
	77.3 %	76.6 %	77.3 %	77.5 %	73.9 %	73 %	77.3 %	76.3 %
Total	1787	1489	2313	1986	1657	1384	1604	12220
	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %

```
# Carga bbdd con comparacion 88 y 08
isco08_88 <- read_excel(here::here("input/data/original/isco08-88.xls"))

isco08_88 <- isco08_88 %>%
  rename(isco08=`ISCO 08 Code`,isco88=`ISCO-88 code`)

# Exploración
valores_repetidos <- table(isco08_88$isco88) # Contar los valores repetidos en la vari
valores_repetidos <- valores_repetidos[valores_repetidos > 1]
valores_repetidos <- table(isco08_88$isco08) # Contar los valores repetidos en la vari
valores_repetidos <- valores_repetidos[valores_repetidos > 1]

# CIUO - ISCO: Recodificación 88->08 4 digitos en datos elsoc
indices <- match(elsoc_long_2016_2023$ciuo88_m03,isco08_88$isco88)
elsoc_long_2016_2023$ciuo08_rec <- isco08_88$isco08[indices]
elsoc_long_2016_2023$ciuo08_rec <- as.numeric(elsoc_long_2016_2023$ciuo08_rec) # es so

# Creación variable isco08 con ciuo08_m03 y ciuo08_rec (ex ciuo88_m03) 4 digitos
elsoc_long_2016_2023$isco08 <- ifelse(!is.na(elsoc_long_2016_2023$ciuo08_rec),elsoc_lo

# Crear una columna con la variable "isco08" adelantada una ola
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
```

```
  group_by(idencuesta) %>%        # Agrupa por id para trabajar en cada individuo
  mutate(isco08_lagged=lag(isco08,n=1)) %>%  # Desplaza isco08 a la siguiente ola
  ungroup()

# Rellenar los valores NA en la variable original
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(isco08=ifelse(!is.na(isco08),isco08,isco08_lagged)) %>%  # Si isco08 es NA, s
  select(-isco08_lagged)                 # Elimina la columna temporal


# CIUO - ISCO: creación de variables con 1 dígito
elsoc_long_2016_2023$isco08_1d <- as.character(elsoc_long_2016_2023$isco08)
elsoc_long_2016_2023$isco08_1d <- substr(elsoc_long_2016_2023$isco08_1d,1,1)
elsoc_long_2016_2023$isco08_1d <- as.numeric(elsoc_long_2016_2023$isco08_1d)

# CIUO - ISCO: creación de variables con 2 dígitos
elsoc_long_2016_2023$isco08_2d <- as.character(elsoc_long_2016_2023$isco08)
elsoc_long_2016_2023$isco08_2d <- substr(elsoc_long_2016_2023$isco08_2d,1,2)
elsoc_long_2016_2023$isco08_2d <- as.numeric(elsoc_long_2016_2023$isco08_2d)

# Tabla de frecuencias y porcentajes 1 dígito
sjt.xtab(elsoc_long_2016_2023$isco08_1d,elsoc_long_2016_2023$ola,
         show.col.prc=TRUE,
         var.labels=c("CIUO08","Ola"),
         show.summary=FALSE,
         title="Frecuencias y porcentajes de Ocupación con 1 dígito, por ola")
```

Table 3: Frecuencias y porcentajes de Ocupación con 1 dígito, por ola

CIUO08	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	
1	89	75	59	47	41	32	48	391
	5 %	5 %	2.5 %	2.3 %	2.5 %	2.3 %	3 %	3.2 %
2	257	198	312	265	250	205	238	1725
	14.4 %	13.3 %	13.3 %	13.2 %	15.1 %	14.8 %	14.8 %	14.1 %
3	362	304	180	150	148	116	202	1462
	20.3 %	20.4 %	7.7 %	7.5 %	8.9 %	8.4 %	12.6 %	11.9 %
4	26	18	187	164	129	107	92	723
	1.5 %	1.2 %	8 %	8.2 %	7.8 %	7.7 %	5.7 %	5.9 %

5	348	279	579	501	383	325	361	2776
	19.5 %	18.8 %	24.8 %	25 %	23.1 %	23.4 %	22.5 %	22.6 %
6	41	37	27	22	30	26	23	206
	2.3 %	2.5 %	1.2 %	1.1 %	1.8 %	1.9 %	1.4 %	1.7 %
7	319	280	346	301	297	259	243	2045
	17.9 %	18.8 %	14.8 %	15 %	17.9 %	18.7 %	15.1 %	16.7 %
8	120	96	202	166	127	103	127	941
	6.7 %	6.5 %	8.6 %	8.3 %	7.6 %	7.4 %	7.9 %	7.7 %
9	220	200	446	391	256	214	272	1999
	12.3 %	13.4 %	19.1 %	19.5 %	15.4 %	15.4 %	16.9 %	16.3 %
Total	1782	1487	2338	2007	1661	1387	1606	12268
	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %

```
# Creación variable Nivel de cualificación
elsoc_long_2016_2023 <- mutate(elsoc_long_2016_2023,
               cualificacion=factor(case_when(isco08_2d<=34&m01>=9~"Experts",
                                              isco08_2d<=34&m01<9~"Skilled",
                                              isco08_2d%in%c(35,41:44,53)~"Skilled",
                                              isco08_2d%in%c(51,54,61,62,71:75,81)&m01
                                              isco08_2d%in%c(51,54,61,62,71:75,81)&m01
                                              isco08_2d%in%c(52,63,82,83,91:99)~"Unski
                                    levels=c("Experts","Skilled","Unskilled")))

# Tabla de frecuencias y porcentajes
sjt.xtab(elsoc_long_2016_2023$cualificacion,elsoc_long_2016_2023$ola,
         show.col.prc=TRUE,
         var.labels=c("Nivel de cualificación","Ola"),
         show.summary=FALSE,
         title="Frecuencias y porcentajes del Nivel de Cualificación, por ola")
```

Table 4: Frecuencias y porcentajes del Nivel de Cualificación, por ola

Nivel de cualificación	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	
Experts	240	181	293	234	244	194	270	1656
	13.5 %	12.2 %	12.5 %	11.7 %	14.7 %	14 %	17 %	13.5 %
Skilled	650	554	679	613	516	419	441	3872
	36.5 %	37.3 %	29.1 %	30.6 %	31.1 %	30.3 %	27.8 %	31.6 %

Unskilled	892	752	1364	1158	900	772	873	6711
	50.1 %	50.6 %	58.4 %	57.8 %	54.2 %	55.7 %	55.1 %	54.8 %
Total	1782	1487	2336	2005	1660	1385	1584	12239
	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %

```
# Crear variable supervisa==1, no supervisa==0
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(supervisa=case_when(m06==0~0,
                             m06>=1~1))


# Crear una columna con la variable "m06" adelantada una ola
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  arrange(ola) %>%  # Ordena por ola, por si acaso no está ordenado
  group_by(idencuesta) %>%        # Agrupa por id para trabajar en cada individuo
  mutate(m06_lagged=lag(m06,n=1)) %>%  # Desplaza "m06" a la siguiente ola
  ungroup()

# Rellenar los valores NA en la variable original
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(m06_full=ifelse(!is.na(m06),m06,m06_lagged)) %>%  # Si m06 es NA, sustituir c
  select(-m06_lagged)

# Crear variable supervisa==1, no supervisa==0
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(supervisa=case_when(m06_full==0~0,
                             m06_full>=1~1))

# Tabla de frecuencias y porcentajes
sjt.xtab(elsoc_long_2016_2023$supervisa,elsoc_long_2016_2023$ola,
         show.col.prc=TRUE,
         var.labels=c("Supervisa","Ola"),
         show.summary=FALSE,
         title="Frecuencias y porcentajes de variable Supervisa, por ola")
```

Table 5: Frecuencias y porcentajes de variable Supervisa, por ola

Supervisa	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	

0	1339	1124	1788	1532	778	643	1112	8316
	74.3 %	74.8 %	77 %	76.9 %	62.2 %	62.7 %	71.1 %	72.6 %
1	463	378	534	461	473	383	452	3144
	25.7 %	25.2 %	23 %	23.1 %	37.8 %	37.3 %	28.9 %	27.4 %
Total	1802	1502	2322	1993	1251	1026	1564	11460
	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %

```r
# Creación de variable posición de clase
elsoc_long_2016_2023 <- mutate(elsoc_long_2016_2023,
                class1=case_when(
                  rel_empleo2=="Empleador"~"Petit bourgeoisie",
                  rel_empleo2=="Autoempleado"&m01>=9~"Petit bourgeoisie",
                  rel_empleo2=="Autoempleado"&m01<9~"Informal",
                  rel_empleo2=="Asalariado"&cualificacion=="Experts"&supervisa==1~"Expe
                  rel_empleo2=="Asalariado"&cualificacion=="Experts"&supervisa==0~"Expe
                  rel_empleo2=="Asalariado"&cualificacion=="Skilled"&supervisa==1~"Supe
                  rel_empleo2=="Asalariado"&cualificacion=="Unskilled"&supervisa==1~"Su
                  rel_empleo2=="Asalariado"&cualificacion=="Skilled"&supervisa==0~"Work
                  rel_empleo2=="Asalariado"&cualificacion=="Unskilled"&supervisa==0~"Wo

elsoc_long_2016_2023$class1 <- factor(elsoc_long_2016_2023$class1,
                      levels=c("Petit bourgeoisie",
                               "Informal",
                               "Expert managers",
                               "Experts",
                               "Supervisors",
                               "Workers"))

# Tabla de frecuencias y porcentajes
sjt.xtab(elsoc_long_2016_2023$class1, elsoc_long_2016_2023$ola,
         show.col.prc = TRUE,
         var.labels = c("Posición de clase 1","Ola"),
         show.summary = FALSE,
         title = "Frecuencias y porcentajes de la Posición de clase 1, por ola")
```

Table 6: Frecuencias y porcentajes de la Posición de clase 1, por ola

Posición de clase 1	Ola							Total
	2016	2017	2018	2019	2021	2022	2023	

Petit bourgeoisie	105 5.9 %	89 6 %	143 6.2 %	121 6.1 %	118 7.6 %	96 7.5 %	84 5.4 %	756 6.4 %
Informal	301 17 %	260 17.6 %	381 16.6 %	323 16.4 %	315 20.4 %	278 21.7 %	280 18 %	2138 18 %
Expert managers	93 5.3 %	70 4.7 %	119 5.2 %	95 4.8 %	115 7.4 %	95 7.4 %	138 8.9 %	725 6.1 %
Experts	112 6.3 %	84 5.7 %	129 5.6 %	103 5.2 %	94 6.1 %	71 5.5 %	99 6.4 %	692 5.8 %
Supervisors	275 15.5 %	229 15.5 %	327 14.2 %	291 14.8 %	275 17.8 %	222 17.3 %	253 16.3 %	1872 15.7 %
Workers	884 49.9 %	743 50.4 %	1199 52.2 %	1039 52.7 %	630 40.7 %	520 40.6 %	700 45 %	5715 48 %
Total	1770 100 %	1475 100 %	2298 100 %	1972 100 %	1547 100 %	1282 100 %	1554 100 %	11898 100 %

```
# Meritocracy

frq(elsoc_long_2016_2023$merit_effort)
```

```
Grado de acuerdo: Las personas son recompensadas por sus esfuerzos (x) <numeric>
# total N=20761 valid N=20629 mean=2.60 sd=1.02

Value |                          Label |    N | Raw % | Valid % | Cum. %
-------------------------------------------------------------------------
    1 |       Totalmente en desacuerdo | 2173 | 10.47 |   10.53 |  10.53
    2 |                  En desacuerdo | 9442 | 45.48 |   45.77 |  56.30
    3 | Ni en desacuerdo ni de acuerdo | 4037 | 19.45 |   19.57 |  75.87
    4 |                     De acuerdo | 4379 | 21.09 |   21.23 |  97.10
    5 |          Totalmente de acuerdo |  598 |  2.88 |    2.90 | 100.00
 <NA> |                          <NA> |  132 |  0.64 |    <NA> |   <NA>
```

```
frq(elsoc_long_2016_2023$merit_talent)
```

```
Grado de acuerdo: Las personas son recompensada por su inteligencia (x) <numeric>
# total N=20761 valid N=20631 mean=2.77 sd=1.03
```

```
Value |                          Label |    N | Raw % | Valid % | Cum. %
--------------------------------------------------------------------------
    1 |       Totalmente en desacuerdo | 1766 |  8.51 |    8.56 |   8.56
    2 |                   En desacuerdo | 8108 | 39.05 |   39.30 |  47.86
    3 | Ni en desacuerdo ni de acuerdo | 4571 | 22.02 |   22.16 |  70.02
    4 |                      De acuerdo | 5577 | 26.86 |   27.03 |  97.05
    5 |          Totalmente de acuerdo |  609 |  2.93 |    2.95 | 100.00
 <NA> |                          <NA> |  130 |  0.63 |    <NA> |   <NA>
```

```r
elsoc_long_2016_2023 <- elsoc_long_2016_2023 %>%
  mutate(
    across(
      .cols = c(merit_effort, merit_talent),
      .fns = ~ car::recode(., recodes = c("1='Strongly disagree'; 2='Disagree';
                                          3='Neither agree nor disagree'; 4='Agree';
                                          5='Strongly agree'"),
                           levels = c("Strongly disagree", "Disagree", "Neither agree
                           as.factor = T)
    )
  )


elsoc_long_2016_2023$merit_effort <- sjlabelled::set_label(elsoc_long_2016_2023$merit_
                         label = "People are rewarded for their efforts")

elsoc_long_2016_2023$merit_talent <- sjlabelled::set_label(elsoc_long_2016_2023$merit_
                         label = "People are rewarded for their intelligence")
```

```r
# Controls

# sex
elsoc_long_2016_2023$sex <- car::recode(elsoc_long_2016_2023$sex,
                           recodes = c("1='Male'; 2='Female'"),
                           levels = c("Male", "Female"),
                           as.factor = T)

elsoc_long_2016_2023$sex <- sjlabelled::set_label(elsoc_long_2016_2023$sex,
                         label = "Gender")

# age
```

```
frq(elsoc_long_2016_2023$age)
```

```
Edad del entrevistado (x) <numeric>
# total N=20761 valid N=20761 mean=48.84 sd=15.43

Value |   N | Raw % | Valid % | Cum. %
--------------------------------------
   18 |  34 |  0.16 |    0.16 |   0.16
   19 |  72 |  0.35 |    0.35 |   0.51
   20 | 113 |  0.54 |    0.54 |   1.05
   21 | 156 |  0.75 |    0.75 |   1.81
   22 | 187 |  0.90 |    0.90 |   2.71
   23 | 231 |  1.11 |    1.11 |   3.82
   24 | 262 |  1.26 |    1.26 |   5.08
   25 | 324 |  1.56 |    1.56 |   6.64
   26 | 318 |  1.53 |    1.53 |   8.17
   27 | 312 |  1.50 |    1.50 |   9.68
   28 | 361 |  1.74 |    1.74 |  11.42
   29 | 345 |  1.66 |    1.66 |  13.08
   30 | 365 |  1.76 |    1.76 |  14.84
   31 | 364 |  1.75 |    1.75 |  16.59
   32 | 369 |  1.78 |    1.78 |  18.37
   33 | 377 |  1.82 |    1.82 |  20.18
   34 | 384 |  1.85 |    1.85 |  22.03
   35 | 373 |  1.80 |    1.80 |  23.83
   36 | 444 |  2.14 |    2.14 |  25.97
   37 | 373 |  1.80 |    1.80 |  27.76
   38 | 411 |  1.98 |    1.98 |  29.74
   39 | 365 |  1.76 |    1.76 |  31.50
   40 | 386 |  1.86 |    1.86 |  33.36
   41 | 391 |  1.88 |    1.88 |  35.24
   42 | 395 |  1.90 |    1.90 |  37.15
   43 | 379 |  1.83 |    1.83 |  38.97
   44 | 361 |  1.74 |    1.74 |  40.71
   45 | 370 |  1.78 |    1.78 |  42.49
   46 | 410 |  1.97 |    1.97 |  44.47
   47 | 411 |  1.98 |    1.98 |  46.45
   48 | 405 |  1.95 |    1.95 |  48.40
   49 | 403 |  1.94 |    1.94 |  50.34
   50 | 451 |  2.17 |    2.17 |  52.51
```

51	417	2.01	2.01	54.52
52	462	2.23	2.23	56.75
53	413	1.99	1.99	58.74
54	463	2.23	2.23	60.97
55	495	2.38	2.38	63.35
56	498	2.40	2.40	65.75
57	449	2.16	2.16	67.91
58	468	2.25	2.25	70.17
59	461	2.22	2.22	72.39
60	481	2.32	2.32	74.70
61	397	1.91	1.91	76.61
62	393	1.89	1.89	78.51
63	386	1.86	1.86	80.37
64	355	1.71	1.71	82.08
65	362	1.74	1.74	83.82
66	316	1.52	1.52	85.34
67	335	1.61	1.61	86.96
68	253	1.22	1.22	88.17
69	260	1.25	1.25	89.43
70	282	1.36	1.36	90.79
71	268	1.29	1.29	92.08
72	230	1.11	1.11	93.18
73	235	1.13	1.13	94.32
74	241	1.16	1.16	95.48
75	245	1.18	1.18	96.66
76	180	0.87	0.87	97.52
77	152	0.73	0.73	98.26
78	105	0.51	0.51	98.76
79	98	0.47	0.47	99.23
80	62	0.30	0.30	99.53
81	40	0.19	0.19	99.73
82	19	0.09	0.09	99.82
83	9	0.04	0.04	99.86
84	11	0.05	0.05	99.91
85	4	0.02	0.02	99.93
86	2	0.01	0.01	99.94
87	2	0.01	0.01	99.95
88	3	0.01	0.01	99.97
89	3	0.01	0.01	99.98
90	2	0.01	0.01	99.99
91	1	0.00	0.00	100.00

```
   92 |    1 |  0.00 |    0.00 | 100.00
 <NA> |    0 |  0.00 |    <NA> |   <NA>
```

```
elsoc_long_2016_2023$age <-
  factor(car::recode(elsoc_long_2016_2023$age,
                     "18:29=1;30:49=2;50:64=3;65:150=4"),
         labels = c('18-29', '30-49', '50-64', '65 or more'))
elsoc_long_2016_2023$age <-
  sjlabelled::set_label(elsoc_long_2016_2023$age,
                        label = c("Age groups"))


# political indentification

frq(elsoc_long_2016_2023$ideo)
```

```
Autoubicacion escala izquierda-derecha (x) <numeric>
# total N=20761 valid N=20443 mean=7.39 sd=3.96

Value |            Label |    N | Raw % | Valid % | Cum. %
-------------------------------------------------------------
    0 |      0 Izquierda | 1146 |  5.52 |    5.61 |   5.61
    1 |                1 |  386 |  1.86 |    1.89 |   7.49
    2 |                2 |  625 |  3.01 |    3.06 |  10.55
    3 |                3 | 1007 |  4.85 |    4.93 |  15.48
    4 |                4 | 1193 |  5.75 |    5.84 |  21.31
    5 |         5 Centro | 5360 | 25.82 |   26.22 |  47.53
    6 |                6 |  749 |  3.61 |    3.66 |  51.20
    7 |                7 |  707 |  3.41 |    3.46 |  54.65
    8 |                8 |  601 |  2.89 |    2.94 |  57.59
    9 |                9 |  177 |  0.85 |    0.87 |  58.46
   10 |       10 Derecha | 1135 |  5.47 |    5.55 |  64.01
   11 | 11 Independiente |  727 |  3.50 |    3.56 |  67.57
   12 |       12 Ninguno | 6630 | 31.93 |   32.43 | 100.00
 <NA> |             <NA> |  318 |  1.53 |    <NA> |   <NA>
```

```
elsoc_long_2016_2023$ideo<-
factor(
  car::recode(
```

```
    elsoc_long_2016_2023$ideo,
    "c(11,12,-888,-999)='Does not identify';c(0,1,2,3,4)='Left';
     c(5)='Center';c(6,7,8,9,10)='Right'"
  ),
  levels = c('Left', 'Center', 'Right', 'Does not identify')
)

elsoc_long_2016_2023$ideo<- factor(elsoc_long_2016_2023$ideo,levels = levels(elsoc_lon

elsoc_long_2016_2023$ideo <-
sjlabelled::set_label(x = elsoc_long_2016_2023$ideo,
                      label = "Political identification")

frq(elsoc_long_2016_2023$ideo)
```

```
Political identification (x) <categorical>
# total N=20761 valid N=20443 mean=2.67 sd=1.17

Value             |    N | Raw % | Valid % | Cum. %
--------------------------------------------------
Left              | 4357 | 20.99 |   21.31 |  21.31
Center            | 5360 | 25.82 |   26.22 |  47.53
Right             | 3369 | 16.23 |   16.48 |  64.01
Does not identify | 7357 | 35.44 |   35.99 | 100.00
<NA>              |  318 |  1.53 |    <NA> |   <NA>
```

```
# Socioeconomic_____________________________________________________________

# Education_____________________________________
elsoc_long_2016_2023$educ <-
  car::recode(elsoc_long_2016_2023$m01,
              "c(1,2,3,4,5,6,7)=1;c(8,9,10)=2; c(-888,-999)=NA")
elsoc_long_2016_2023$educ <-
  factor(elsoc_long_2016_2023$educ,
         labels = c("Less than Universitary","Universitary"))

#reverse education, reference level is the highest level
#elsoc_long_2016_2023$educ <- forcats::fct_rev(elsoc_long_2016_2023$educ)
```

```r
elsoc_long_2016_2023$educ <-
sjlabelled::set_label(x = elsoc_long_2016_2023$educ,
                      label = "Education")
sjmisc::frq(elsoc_long_2016_2023$educ)
```

```
Education (x) <categorical>
# total N=20761 valid N=20746 mean=1.19 sd=0.39

Value                  |     N | Raw % | Valid % | Cum. %
-----------------------------------------------------------
Less than Universitary | 16786 | 80.85 |   80.91 |  80.91
Universitary           |  3960 | 19.07 |   19.09 | 100.00
<NA>                   |    15 |  0.07 |    <NA> |   <NA>
```

```r
#Recoding of education to years based on casen 2017.
elsoc_long_2016_2023$educyear<- as.numeric(
  car::recode(elsoc_long_2016_2023$m01,
              "1=0;2=4.3;3=7.5;4=9.8;5=12.02;6=13.9;
               7=14.8;8=14.9;9=16.9;10=19.07;c(-888,-999)=NA",
              as.numeric = T))

elsoc_long_2016_2023$educyear <-
sjlabelled::set_label(x = elsoc_long_2016_2023$educyear,
                      label = "Education in years")

class(elsoc_long_2016_2023$educyear)
```

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

```r
sjmisc::frq(elsoc_long_2016_2023$educyear)
```

```
Education in years (x) <numeric>
# total N=20761 valid N=20746 mean=11.56 sd=3.99

Value |    N | Raw % | Valid % | Cum. %
---------------------------------------
 0.00 |  200 |  0.96 |    0.96 |   0.96
 4.30 | 2412 | 11.62 |   11.63 |  12.59
```

```
 7.50 | 1992 |  9.59 |    9.60 |  22.19
 9.80 | 2692 | 12.97 |   12.98 |  35.17
12.02 | 6122 | 29.49 |   29.51 |  64.68
13.90 |  749 |  3.61 |    3.61 |  68.29
14.80 | 2619 | 12.61 |   12.62 |  80.91
14.90 | 1197 |  5.77 |    5.77 |  86.68
16.90 | 2413 | 11.62 |   11.63 |  98.31
19.07 |  350 |  1.69 |    1.69 | 100.00
 <NA> |   15 |  0.07 |    <NA> |   <NA>
```

```r
# Reshape long to wide

df_study1_long <- elsoc_long_2016_2023 %>%
  select(idencuesta,
         ola,
         muestra,
         ponderador_long_total,
         segmento,
         estrato,
         just_pension,
         class1,
         merit_effort,
         merit_talent,
         educ,
         educyear,
         sex,
         age,
         ideo)

df_study1_wide <- df_study1_long %>%
  tidyr::pivot_wider(id_cols = c("idencuesta","muestra"),
                     names_from = "ola",
                     values_from = names(select(df_study1_long,ponderador_long_total,s

# fix data to w01 values
df_study1_wide$class1_2 <-df_study1_wide$class1_1 #class
df_study1_wide$class1_3 <-df_study1_wide$class1_1
df_study1_wide$class1_4 <-df_study1_wide$class1_1
df_study1_wide$class1_5 <-df_study1_wide$class1_1
df_study1_wide$class1_6 <-df_study1_wide$class1_1
df_study1_wide$class1_7 <-df_study1_wide$class1_1
```

```r
df_study1_wide$age_2 <-df_study1_wide$age_1 #age
df_study1_wide$age_3 <-df_study1_wide$age_1
df_study1_wide$age_4 <-df_study1_wide$age_1
df_study1_wide$age_5 <-df_study1_wide$age_1
df_study1_wide$age_6 <-df_study1_wide$age_1
df_study1_wide$age_7 <-df_study1_wide$age_1

df_study1_wide$sex_2 <-df_study1_wide$sex_1 #sex
df_study1_wide$sex_3 <-df_study1_wide$sex_1
df_study1_wide$sex_4 <-df_study1_wide$sex_1
df_study1_wide$sex_5 <-df_study1_wide$sex_1
df_study1_wide$sex_6 <-df_study1_wide$sex_1
df_study1_wide$sex_7 <-df_study1_wide$sex_1

df_study1_wide$educ_2 <-df_study1_wide$educ_1 #education
df_study1_wide$educ_3 <-df_study1_wide$educ_1
df_study1_wide$educ_4 <-df_study1_wide$educ_1
df_study1_wide$educ_5 <-df_study1_wide$educ_1
df_study1_wide$educ_6 <-df_study1_wide$educ_1
df_study1_wide$educ_7 <-df_study1_wide$educ_1

df_study1_wide$educyear_2 <-df_study1_wide$educyear_1 #education years
df_study1_wide$educyear_3 <-df_study1_wide$educyear_1
df_study1_wide$educyear_4 <-df_study1_wide$educyear_1
df_study1_wide$educyear_5 <-df_study1_wide$educyear_1
df_study1_wide$educyear_6 <-df_study1_wide$educyear_1
df_study1_wide$educyear_7 <-df_study1_wide$educyear_1


df_study1_wide$ideo_2 <-df_study1_wide$ideo_1 # political position
df_study1_wide$ideo_3 <-df_study1_wide$ideo_1
df_study1_wide$ideo_4 <-df_study1_wide$ideo_1
df_study1_wide$ideo_5 <-df_study1_wide$ideo_1
df_study1_wide$ideo_6 <-df_study1_wide$ideo_1
df_study1_wide$ideo_7 <-df_study1_wide$ideo_1


dim(df_study1_wide)
```

```
[1] 4447   86
```

```
sjPlot::view_df(df_study1_wide,
                show.frq = T,show.values = T,show.na = T,show.prc = T, show.type = T)
```

```
just_pension_5 [28]
```

Table 7: Data frame: df_study1_wide

ID	Name	Type	Label	missings	Values	Value Labels	Freq.	%
1	idencuesta	numeric	Folio identificador de participante	0 (0.00%)	*range: 1101011-NA*			
2	muestra	numeric	Identificador de muestra de encuesta	0 (0.00%)	1 2	Muestra Original Muestra Refresco	2928 1519	65.84 34.16
3	ponderador_longitudinal_1	numeric	Total ponderador longitudinal (2016-2022). Muestra completa	1520 (34.18%)	*range: 0.0-6.6*			
4	ponderador_longitudinal_2	numeric	Total ponderador longitudinal (2016-2022). Muestra completa	1974 (44.39%)	*range: 0.0-7.7*			

5	ponderador_longitudinal	numeric	Total de casos longitudinal (2016-2022). Muestra completa	Ponderador 3	699 (15.72%)	*range: 0.0-9.0*
6	ponderador_longitudinal	numeric	Total de casos longitudinal (2016-2022). Muestra completa	Ponderador 4	1030 (23.16%)	*range: 0.0-11.8*
7	ponderador_longitudinal	numeric	Total de casos longitudinal (2016-2022). Muestra completa	Ponderador 5	1707 (38.39%)	*range: 0.0-13.8*
8	ponderador_longitudinal	numeric	Total de casos longitudinal (2016-2022). Muestra completa	Ponderador 6	1717 (38.61%)	*range: 0.0-15.0*
9	ponderador_longitudinal	numeric	Total de casos longitudinal (2016-2022). Muestra completa	Ponderador 7	1721 (38.70%)	*range: 0.0-16.4*

10	segmento_1	numeric	Segmento	1520 (34.18%)	*range: 110101-1420411*			
11	segmento_2	numeric	Segmento	1974 (44.39%)	*range: 110101-1420411*			
12	segmento_3	numeric	Segmento	699 (15.72%)	*range: 110101-1420411*			
13	segmento_4	numeric	Segmento	1030 (23.16%)	*range: 110101-1420411*			
14	segmento_5	numeric	Segmento	1707 (38.39%)	*range: 110101-1420411*			
15	segmento_6	numeric	Segmento	1717 (38.61%)	*range: 110101-1420411*			
16	segmento_7	numeric	Segmento	1721 (38.70%)	*range: 110101-1420411*			
17	estrato_1	numeric	Estrato muestral	1520 (34.18%)	1 2 3 4 5 6	Gran Santiago Gran Valparaiso Gran Concepcion Ciudades Grandes Ciudades Medianas Ciudades Pequennias	720 375 391 408 567 466	24.60 12.81 13.36 13.94 19.37 15.92

18	estrato_2	numeric	Estrato muestral	1974 (44.39%)	1 Gran Santiago	597	24.14
					2 Gran Valparaiso	314	12.70
					3 Gran Concepcion	335	13.55
					4 Ciudades Grandes	363	14.68
					5 Ciudades Medianas	457	18.48
					6 Ciudades Pequennias	407	16.46
19	estrato_3	numeric	Estrato muestral	699 (15.72%)	1 Gran Santiago	934	24.92
					2 Gran Valparaiso	437	11.66
					3 Gran Concepcion	514	13.71
					4 Ciudades Grandes	623	16.62
					5 Ciudades Medianas	636	16.97
					6 Ciudades Pequennias	604	16.12

20	estrato_4	numeric	Estrato muestral	1030 (23.16%)	1 Gran Santiago 2 Gran Valparaiso 3 Gran Concepcion 4 Ciudades Grandes 5 Ciudades Medianas 6 Ciudades Pequennias	833 395 467 568 593 561	24.38 11.56 13.67 16.62 17.35 16.42
21	estrato_5	numeric	Estrato muestral	1707 (38.39%)	1 Gran Santiago 2 Gran Valparaiso 3 Gran Concepcion 4 Ciudades Grandes 5 Ciudades Medianas 6 Ciudades Pequennias	671 316 394 456 470 433	24.49 11.53 14.38 16.64 17.15 15.80

22	estrato_6	numeric	Estrato muestral	1717 (38.61%)	1 Gran Santiago 2 Gran Valparaiso 3 Gran Concepcion 4 Ciudades Grandes 5 Ciudades Medianas 6 Ciudades Pequennias	688 298 369 449 440 486	25.20 10.92 13.52 16.45 16.12 17.80
23	estrato_7	numeric	Estrato muestral	1721 (38.70%)	1 Gran Santiago 2 Gran Valparaiso 3 Gran Concepcion 4 Ciudades Grandes 5 Ciudades Medianas 6 Ciudades Pequennias	665 308 359 441 467 486	24.39 11.30 13.17 16.18 17.13 17.83

24	just_pension_categorical_1	Pension distributive justice	1524 (34.27%)		Strongly disagree	862	29.49
					disagree	1337	45.74
					Disagree	228	7.80
					Neither agree nor disagree	443	15.16
					Agree	53	1.81
					Strongly agree		
25	just_pension_categorical_2	Pension distributive justice	1982 (44.57%)		Strongly disagree	906	36.75
					disagree	1022	41.46
					Disagree	172	6.98
					Neither agree nor disagree	314	12.74
					Agree	51	2.07
					Strongly agree		
26	just_pension_categorical_3	Pension distributive justice	718 (16.15%)		Strongly disagree	1253	33.60
					disagree	1517	40.68
					Disagree	283	7.59
					Neither agree nor disagree	573	15.37
					Agree	103	2.76
					Strongly agree		

27	just_pension_categorical_4	Pension distributive justice	1038 (23.34%)		Strongly disagree	963	28.25
					disagree	1598	46.88
					Disagree	321	9.42
					Neither agree nor disagree	467	13.70
					Agree	60	1.76
					Strongly agree		
29	just_pension_categorical_6	Pension distributive justice	1726 (38.81%)		Strongly disagree	415	15.25
					disagree	1207	44.36
					Disagree	348	12.79
					Neither agree nor disagree	661	24.29
					Agree	90	3.31
					Strongly agree		
30	just_pension_categorical_7	Pension distributive justice	1728 (38.86%)		Strongly disagree	490	18.02
					disagree	1121	41.23
					Disagree	388	14.27
					Neither agree nor disagree	629	23.13
					Agree	91	3.35
					Strongly agree		

31	class1_1	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
					301	17.01
					93	5.25
				Informal	112	6.33
				Expert managers	275	15.54
					884	49.94
				Experts		
				Supervisors		
				Workers		
32	class1_2	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
					301	17.01
					93	5.25
				Informal	112	6.33
				Expert managers	275	15.54
					884	49.94
				Experts		
				Supervisors		
				Workers		
33	class1_3	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
					301	17.01
					93	5.25
				Informal	112	6.33
				Expert managers	275	15.54
					884	49.94
				Experts		
				Supervisors		
				Workers		
34	class1_4	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
					301	17.01
					93	5.25
				Informal	112	6.33
				Expert managers	275	15.54
					884	49.94
				Experts		
				Supervisors		
				Workers		

35	class1_5	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
				Informal	301	17.01
				Expert managers	93	5.25
				Experts	112	6.33
				Supervisors	275	15.54
				Workers	884	49.94
36	class1_6	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
				Informal	301	17.01
				Expert managers	93	5.25
				Experts	112	6.33
				Supervisors	275	15.54
				Workers	884	49.94
37	class1_7	categorical	2677 (60.20%)	Petit bourgeoisie	105	5.93
				Informal	301	17.01
				Expert managers	93	5.25
				Experts	112	6.33
				Supervisors	275	15.54
				Workers	884	49.94

38	merit_effort_categorical	People are rewarded for their efforts	1538 (34.59%)	Strongly disagree	357	12.27
				Disagree	1331	45.75
				Neither agree nor disagree	497	17.08
				Agree	646	22.21
				Strongly agree	78	2.68
39	merit_effort_categorical	People are rewarded for their efforts	1988 (44.70%)	Strongly disagree	282	11.47
				Disagree	1057	42.98
				Neither agree nor disagree	478	19.44
				Agree	556	22.61
				Strongly agree	86	3.50
40	merit_effort_categorical	People are rewarded for their efforts	737 (16.57%)	Strongly disagree	347	9.35
				Disagree	1482	39.95
				Neither agree nor disagree	797	21.48
				Agree	925	24.93
				Strongly agree	159	4.29

41	merit_effort_categorical	People are rewarded for their efforts	1046 (23.52%)	Strongly disagree	415	12.20
				disagree	1604	47.16
				Disagree	631	18.55
				Neither agree nor disagree	653	19.20
				Agree	98	2.88
				Strongly agree		
42	merit_effort_categorical	People are rewarded for their efforts	1735 (39.02%)	Strongly disagree	237	8.74
				disagree	1415	52.18
				Disagree	422	15.56
				Neither agree nor disagree	567	20.91
				Agree	71	2.62
				Strongly agree		
43	merit_effort_categorical	People are rewarded for their efforts	1729 (38.88%)	Strongly disagree	265	9.75
				disagree	1323	48.68
				Disagree	587	21.60
				Neither agree nor disagree	491	18.06
				Agree	52	1.91
				Strongly agree		

44	merit_effort_cat7	categorical	People are rewarded for their efforts	1727 (38.84%)	Strongly disagree disagree Disagree Neither agree nor disagree Agree Strongly agree	270 1230 625 541 54	9.93 45.22 22.98 19.89 1.99
45	merit_talent_cat1	categorical	People are rewarded for their intelligence	1540 (34.63%)	Strongly disagree disagree Disagree Neither agree nor disagree Agree Strongly agree	288 1163 559 814 83	9.91 40.01 19.23 28.00 2.86
46	merit_talent_cat2	categorical	People are rewarded for their intelligence	1987 (44.68%)	Strongly disagree disagree Disagree Neither agree nor disagree Agree Strongly agree	228 901 555 675 101	9.27 36.63 22.56 27.44 4.11

47	merit_talent_3	categorical	People are rewarded for their intelligence	739 (16.62%)	Strongly disagree Disagree Neither agree nor disagree Agree Strongly agree	277 1257 838 1177 159	7.47 33.90 22.60 31.74 4.29
48	merit_talent_4	categorical	People are rewarded for their intelligence	1047 (23.54%)	Strongly disagree Disagree Neither agree nor disagree Agree Strongly agree	340 1329 775 862 94	10.00 39.09 22.79 25.35 2.76
49	merit_talent_5	categorical	People are rewarded for their intelligence	1733 (38.97%)	Strongly disagree Disagree Neither agree nor disagree Agree Strongly agree	168 1290 495 694 67	6.19 47.53 18.24 25.57 2.47

50	merit_talent_6	categorical	People are rewarded for their intelligence	1726 (38.81%)	Strongly disagree	238	8.75
					dis-agree	1108	40.72
					Disagree	667	24.51
					Neither agree nor disagree	658	24.18
					Agree	50	1.84
					Strongly agree		
51	merit_talent_7	categorical	People are rewarded for their intelligence	1726 (38.81%)	Strongly disagree	227	8.34
					dis-agree	1060	38.96
					Disagree	682	25.06
					Neither agree nor disagree	697	25.62
					Agree	55	2.02
					Strongly agree		
52	educ_1	categorical	Education	1522 (34.23%)	Less than Universitary	2390	81.71
					Universitary	535	18.29
53	educ_2	categorical	Education	1522 (34.23%)	Less than Universitary	2390	81.71
					Universitary	535	18.29
54	educ_3	categorical	Education	1522 (34.23%)	Less than Universitary	2390	81.71
					Universitary	535	18.29

55	educ_4	categorical	Education	1522 (34.23%)		Less than Universitary	2390	81.71
						Universitary	535	18.29
56	educ_5	categorical	Education	1522 (34.23%)		Less than Universitary	2390	81.71
						Universitary	535	18.29
57	educ_6	categorical	Education	1522 (34.23%)		Less than Universitary	2390	81.71
						Universitary	535	18.29
58	educ_7	categorical	Education	1522 (34.23%)		Less than Universitary	2390	81.71
						Universitary	535	18.29
59	educyear_1	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
60	educyear_2	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
61	educyear_3	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
62	educyear_4	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
63	educyear_5	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
64	educyear_6	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			

65	educyear_n	numeric	Education in years	1522 (34.23%)	*range: 0.0-19.1*			
66	sex_1	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
67	sex_2	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
68	sex_3	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
69	sex_4	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
70	sex_5	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
71	sex_6	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
72	sex_7	categorical	Gender	1520 (34.18%)		Male	1163	39.73
						Female	1764	60.27
73	age_1	categorical	Age groups	1520 (34.18%)		18-29	506	17.29
						30-49	1157	39.53
						50-64	839	28.66
						65 or more	425	14.52
74	age_2	categorical	Age groups	1520 (34.18%)		18-29	506	17.29
						30-49	1157	39.53
						50-64	839	28.66
						65 or more	425	14.52
75	age_3	categorical	Age groups	1520 (34.18%)		18-29	506	17.29
						30-49	1157	39.53
						50-64	839	28.66
						65 or more	425	14.52
76	age_4	categorical	Age groups	1520 (34.18%)		18-29	506	17.29
						30-49	1157	39.53
						50-64	839	28.66
						65 or more	425	14.52

77	age_5	categorical	Age groups	1520 (34.18%)	18-29	506	17.29
					30-49	1157	39.53
					50-64	839	28.66
					65 or more	425	14.52
78	age_6	categorical	Age groups	1520 (34.18%)	18-29	506	17.29
					30-49	1157	39.53
					50-64	839	28.66
					65 or more	425	14.52
79	age_7	categorical	Age groups	1520 (34.18%)	18-29	506	17.29
					30-49	1157	39.53
					50-64	839	28.66
					65 or more	425	14.52
80	ideo_1	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22
81	ideo_2	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22
82	ideo_3	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22
83	ideo_4	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22

84	ideo_5	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22
85	ideo_6	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22
86	ideo_7	categorical	Political identification	1586 (35.66%)	Left	583	20.38
					Center	604	21.11
					Right	409	14.30
					Does not identify	1265	44.22

```
# reshape from long to wide
pacman::p_load(datasets,data.table)
df_study1_long <- data.table::melt.data.table(data.table::setDT(df_study1_wide),
              id.vars = c("idencuesta","muestra"),
              variable.name = c("ola"),
              measure =
patterns("^ponderador_long_total_", "^segmento_", "^estrato_", "^just_pension_","^clas
              value.name = c("ponderador_long_total","segmento","estrato", "just_pensi
              na.rm = F,value.factor = TRUE
              )
```

```
names(df_study1_long) #check names of long dataset
```

```
 [1] "idencuesta"            "muestra"               "ola"
 [4] "ponderador_long_total" "segmento"              "estrato"
 [7] "just_pension"          "class"                 "merit_effort"
[10] "merit_talent"          "educ"                  "educyear"
[13] "sex"                   "age"                   "ideo"
```

```r
dim(df_study1_long) #check dimensions of the dataframe
```

```
[1] 31129    15
```

```r
# Original dataset with 7 waves
df_study2_long <- df_study1_long

# filter the dataset for the waves 1 to 4 and 6 to 7
df_study1_long <-
df_study1_long %>%
  filter(ola %in% c(1,2,3,4,6,7)) %>%
  mutate(ola=factor(ola,levels = 1:7,labels = 1:7))
dim(df_study1_long) #check, now is OK
```

```
[1] 26682    15
```

```r
# df_study1_long <-
# set_label(x = df_study1_long,
#           label = get_label(select(df_study1,names(df_study1_long))))


#_______________________________________________________________________________
# obtain the idencuesta for wave 7
ids <-
  elsoc_long_2016_2023 %>%
  select(idencuesta,ola) %>%
  filter(ola==7) %>%
  sjmisc::frq(idencuesta,show.na = F) %>% as.data.frame()


# filter data by the idencuesta of t7
df_study1_long_t7 <-
  df_study1_long %>%
  filter(idencuesta %in% ids$val)

names(df_study1_long_t7)
```

```
 [1] "idencuesta"            "muestra"               "ola"
```

```
 [4] "ponderador_long_total" "segmento"              "estrato"
 [7] "just_pension"          "class"                 "merit_effort"
[10] "merit_talent"          "educ"                  "educyear"
[13] "sex"                   "age"                   "ideo"
```

```
dim(df_study1_long_t7)
```

```
[1] 16356    15
```

```
sjmisc::frq(df_study1_long_t7$ola)
```

```
x <categorical>
# total N=16356 valid N=16356 mean=3.83 sd=2.11

Value |    N | Raw % | Valid % | Cum. %
--------------------------------------
    1 | 2726 | 16.67 |   16.67 |  16.67
    2 | 2726 | 16.67 |   16.67 |  33.33
    3 | 2726 | 16.67 |   16.67 |  50.00
    4 | 2726 | 16.67 |   16.67 |  66.67
    5 |    0 |  0.00 |    0.00 |  66.67
    6 | 2726 | 16.67 |   16.67 |  83.33
    7 | 2726 | 16.67 |   16.67 | 100.00
 <NA> |    0 |  0.00 |    <NA> |   <NA>
```

```
save(df_study1_long,file = here::here("input/data/proc/df_study1_long.RData"))
save(df_study1_long_t7,file = here::here("input/data/proc/df_study1_long_t7.RData"))
save(df_study2_long,file = here::here("input/data/proc/df_study2_long.RData"))
```