

descriptivos-meritocracia

Equipo EDUMER

29-11-2021

Contents

merge bases	2
Dependientes	4
independientes	6
School	8
graficos descriptivos	10
dependientes	10
independientes	11

```
pacman::p_load(haven, dplyr, summarytools, sjmisc, car, sjlabelled, webshot, sjPlot, ggplot2, corrplot,
load("input/data/proc/alumnos_patterns.RData")
load("input/data/proc/apoderados_patterns.RData")
```

```
data_est <- alumnos_patterns %>% dplyr::select(idalumno,
                                                sexo,
                                                codigoCurso,
                                                mrbd,
                                                cod_depe2, #dependencia administrativa
                                                cod_grupo, #categorizacion socioeconomica
                                                simce_lect=prom_lect8b_rbd, #Promedio lectura escuela
                                                simce_mate=prom_mate8b_rbd, # Promedio matematica escuela
                                                inteligencia_esc=est_p6_1, # inteligencia escolar
                                                esfuerzo_esc=est_p6_2, # esfuerzo escolar
                                                esfuerzo_soc=est_p38_2, #esfuerzo social
                                                merito_soc=est_p38_3, #merito social
                                                inteligencia_soc=est_p38_4, #inteligencia social
                                                just_pension=est_p39_1, #justicia pension
                                                just_educ=est_p39_2, # justicia educacion
                                                just_salud=est_p39_3, # justicia salud
                                                redistribucion=est_p39_4, # redistribucion
                                                libros=est_p3, # cantidad de libros
                                                pc=est_p4_1, # tiene pc
                                                tablet=est_p4_2, # tiene tablet
                                                celular=est_p4_3 # tiene celular
                                                ) %>% as.data.frame(.)

data_apod <- apoderados_patterns %>% dplyr::select(idalumno,
                                                    educacion=educ_max, # nivel educacional mas alto
```

```
internet=apod_p9 # conexion a internet en la casa
) %>% as.data.frame())
```

merge bases

```
data <- merge(data_est, data_apod, by="idalumno")
rm(apoderados_patterns, alumnos_patterns, data_est, data_apod)

# Educacion. 10 categorías, se recodifican en 4 y una de no sabe y no responde
data$educacion_rec <- ifelse(is.na(data$educacion), "Ns/Nr", data$educacion)
data <- data %>% rowwise() %>% mutate(educacion_rec = case_when(educacion_rec==1~"8th grade or less",
                                                                educacion_rec==2~"8th grade or less",
                                                                educacion_rec==3~"8th grade or less",
                                                                educacion_rec==4~"Secondary Education",
                                                                educacion_rec==5~"Secondary Education",
                                                                educacion_rec==6~"Higher tec. education",
                                                                educacion_rec==7~"Higher tec. education",
                                                                educacion_rec==8~"University or Postgraduate",
                                                                educacion_rec==9~"University or Postgraduate",
                                                                educacion_rec==10~"University or Postgraduate",
                                                                educacion_rec=="Ns/Nr"~"Missing"
                                                                ))
data$educacion_rec <- factor(data$educacion_rec, levels = c("8th grade or less", "Secondary Education",
data$educacion_rec <- set_label(data$educacion_rec, label = "Parental educational level")

## Proporción de Universidad o posgrado por escuela
data <- data %>% group_by(mrbd) %>%
  mutate(mean_educ = mean(educacion_rec=="University or Postgraduate", na.rm=T))
data$mean_educ <- set_label(data$mean_educ, label = "Proportion of parents with university level by school")

# Internet
data <- data %>% set_na(internet, na = c(0,99), drop.levels = FALSE, as.tag = FALSE) # recode missings

data$internet_rec <- ifelse(is.na(data$internet), "Missing", data$internet)
data$internet_rec <- factor(data$internet_rec, levels = c("1", "2"), labels = c("Yes", "No"))
# libros
data$libros_rec <- as.numeric(data$libros)
data <- data %>% rowwise() %>% mutate(libros_rec = case_when(libros_rec==1 ~ "Less than 25",
                                                            libros_rec==2 ~ "Less than 25",
                                                            libros_rec==3 ~ "More than 25",
                                                            libros_rec==4 ~ "More than 25",
                                                            libros_rec==5 ~ "More than 25"))
data$libros_rec <- factor(data$libros_rec, levels = c("Less than 25", "More than 25"))
data$libros_rec <- set_label(data$libros_rec, label = "Number of books at home")
## Promedio simce
data <- data %>% rowwise() %>% mutate(prom_simce = mean(simce_lect, simce_mate))
summary(data$prom_simce)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      180.0   231.0   247.0   246.5   262.0   304.0

data$simce <- ntile(data$prom_simce, 3)
table(data$simce)
```

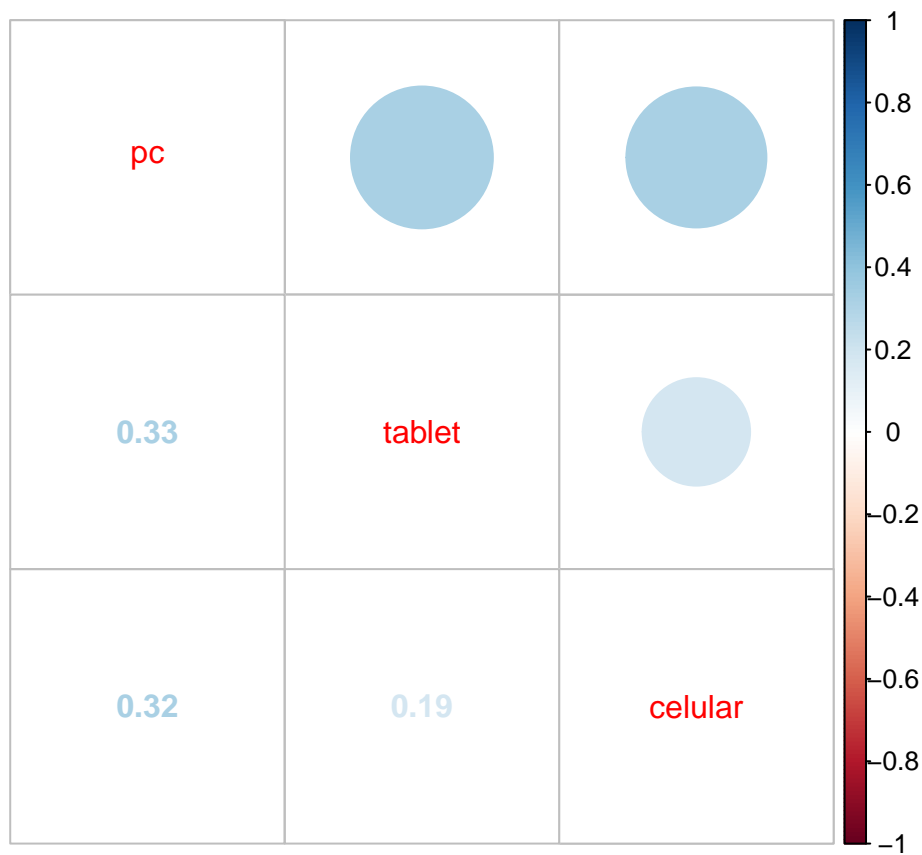
```
##
##      1      2      3
## 2091 2091 2090
```

```
data$simce <- factor(data$simce, labels = c("Low", "Medium", "High"))
data$simce <- set_label(data$simce, label = "SIMCE score by school")
```

```
# Acceso a tecnologia
```

```
data$pc <- as.numeric(data$pc)
data$tablet <- as.numeric(data$tablet)
data$celular <- as.numeric(data$celular)
```

```
corrplot.mixed(cor(dplyr::select(data, pc, tablet, celular),
                  method = "spearman",
                  use = "complete.obs"))
```



```
ltm::cronbach.alpha(data %>%
  dplyr::select(pc, tablet, celular), na.rm=TRUE)
```

```
##
## Cronbach's alpha for the 'data %>% dplyr::select(pc, tablet, celular)' data-set
##
## Items: 3
## Sample units: 6272
## alpha: 0.528
```

```
data <- data %>% rowwise() %>% dplyr::mutate(acc_tec = sum(pc, tablet, celular, na.rm = T))
summary(data$acc_tec)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000   6.000   8.000   7.775   9.000  12.000
```

```
data$acc_tec <- set_label(data$acc_tec,label = "Technology access index")
```

Dependientes

```
dependientes <- data %>% dplyr::select(just_pension, just_educ, just_salud)
```

```
dependientes$pension_rec <- as.numeric(dependientes$just_pension)
```

```
dependientes$educ_rec <- as.numeric(dependientes$just_salud)
```

```
dependientes$salud_rec <- as.numeric(dependientes$just_educ)
```

```
#dependientes <- dependientes %>% dplyr::mutate(pension_rec = 5-pension_rec,
#                                              educ_rec = 5-educ_rec,
#                                              salud_rec = 5-salud_rec)
```

```
ltm::cronbach.alpha(dependientes %>%
  dplyr::select(educ_rec, salud_rec, pension_rec), na.rm=TRUE)
```

```
##
```

```
## Cronbach's alpha for the 'dependientes %>% dplyr::select(educ_rec, salud_rec, pension_rec)' data-set
```

```
##
```

```
## Items: 3
```

```
## Sample units: 6272
```

```
## alpha: 0.867
```

```
dependientes <- dependientes %>% rowwise() %>% dplyr::mutate(bienestar = mean(c(educ_rec, salud_rec, pension_rec)))
dependientes$bienestar <- round(dependientes$bienestar, 2)
```

```
dependientes <- dependientes %>% dplyr::select(-pension_rec, -educ_rec, -salud_rec)
```

```
dependientes$just_pension <- factor(dependientes$just_pension, levels = c("Muy desacuerdo", "Desacuerdo", "De acuerdo", "Muy de acuerdo"))
```

```
dependientes$just_pension <- set_label(dependientes$just_pension, label = "It is just that in Chile people who")
```

```
dependientes$just_salud <- factor(dependientes$just_salud, levels = c("Muy desacuerdo", "Desacuerdo", "De acuerdo", "Muy de acuerdo"))
```

```
dependientes$just_salud <- set_label(dependientes$just_salud,label = "It is just that in Chile people who")
```

```
dependientes$just_educ <- factor(dependientes$just_educ, levels = c("Muy desacuerdo", "Desacuerdo", "De acuerdo", "Muy de acuerdo"))
```

```
dependientes$just_educ <- set_label(dependientes$just_educ,label = "It is just that in Chile people who")
```

```
dependientes$bienestar <- set_label(dependientes$bienestar,label = "Inequality justification index")
```

```
#dependientes$redistribucion <- factor(dependientes$redistribucion, levels = c("Muy desacuerdo", "Desacuerdo", "De acuerdo", "Muy de acuerdo"))
```

```
#dependientes$redistribucion <- set_label(dependientes$redistribucion,label = "The Chilean government's redistribution policy")
```

```
df<-dfSummary(dependientes,
  plain.ascii = TRUE,
  style = "multiline",
  tmp.img.dir = "/tmp",
  graph.magnif = 0.75,
  headings = F, # encabezado
  varnumbers = F, # num variable
  labels.col = T, # etiquetas
  na.col = F, # missing
  graph.col = F, # plot
```

```
valid.col = T, # n valido
col.widths = c(30,10,10,10))
```

```
df$Variable <- NULL # delete variable column
df
```

Label	Stats / Values	Freqs (% of Valid)	Valid
It is just that in Chile people with higher incomes can have better pensions than people with low incomes	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	1837 (30.6%) 1945 (32.4%) 1622 (27.0%) 608 (10.1%)	6012 (95.9%)
It is just that in Chile people who can pay have a better education for their children	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	1766 (29.7%) 1732 (29.1%) 1704 (28.6%) 750 (12.6%)	5952 (94.9%)
It is just that in Chile people with higher incomes can access better health services than people with low incomes	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	2254 (38.0%) 1685 (28.4%) 1401 (23.6%) 593 (10.0%)	5933 (94.6%)
Inequality justification index	Mean (sd) : 2.2 (0.9) min < med < max: 1 < 2 < 4 IQR (CV) : 1.7 (0.4)	13 distinct values	6077 (96.9%)

```
view(df, footnote=NA, file = "output/tables/desc01.html")
```

```
## Output file written: C:\Users\kevin\OneDrive\Documentos\meritocracia-escuela-agencia\output\tables\d
webshot::webshot(url = "output/tables/desc01.html" ,file = "output/tables/desc01.png")
```

Label	Stats / Values	Freqs (% of Valid)	Valid
It is just that in Chile people with higher incomes can have better pensions than people with low incomes	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	1837 (30.6%) 1945 (32.4%) 1622 (27.0%) 608 (10.1%)	6012 (95.9%)
It is just that in Chile people who can pay have a better education for their children	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	1766 (29.7%) 1732 (29.1%) 1704 (28.6%) 750 (12.6%)	5952 (94.9%)
It is just that in Chile people with higher incomes can access better health services than people with low incomes	1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree	2254 (38.0%) 1685 (28.4%) 1401 (23.6%) 593 (10.0%)	5933 (94.6%)
Inequality justification index	Mean (sd) : 2.2 (0.9) min ≤ med ≤ max: 1 ≤ 2 ≤ 4 IQR (CV) : 1.7 (0.4)	13 distinct values	6077 (96.9%)

independientes

```

independientes <- data %>% dplyr::select(inteligencia_esc, esfuerzo_esc, inteligencia_soc, esfuerzo_soc

independientes$inteligencia_esc <- factor(independientes$inteligencia_esc, levels = c("Muy desacuerdo",
independientes$inteligencia_esc <- set_label(independientes$inteligencia_esc, label = "Intelligence is

independientes$esfuerzo_esc <- factor(independientes$esfuerzo_esc, levels = c("Muy desacuerdo", "Desacu
independientes$esfuerzo_esc <- set_label(independientes$esfuerzo_esc, label = "Effort is important to g

independientes$esfuerzo_soc <- factor(independientes$esfuerzo_soc, levels = c("Muy desacuerdo", "Desacu
independientes$esfuerzo_soc <- set_label(independientes$esfuerzo_soc, label = "In Chile, people are rew

independientes$merito_soc <- factor(independientes$merito_soc, levels = c("Muy desacuerdo", "Desacuerdo
independientes$merito_soc <- set_label(independientes$merito_soc, label = "In Chile, people get what th

independientes$inteligencia_soc <- factor(independientes$inteligencia_soc, levels = c("Muy desacuerdo",
independientes$inteligencia_soc <- set_label(independientes$inteligencia_soc, label = "In Chile, people

df2<- dfSummary(independientes,
  plain.ascii = FALSE,

```

```

style = "grid",
tmp.img.dir = "/tmp",
graph.magnif = 0.75,
headings = F, # encabezado
varnumbers = F, # num variable
labels.col = T, # etiquetas
na.col = F, # missing
graph.col = F, # plot
valid.col = T, # n valido
col.widths = c(20,30,10,10,10))

```

```

df2$Variable <- c("School talent", "School effort", "Social talent", "Social effort", "Deservingness",
df2

```

```

##
## +-----+-----+-----+-----+
## | Variable      | Label                                     | Stats / Values          | Freqs (N) |
## +-----+-----+-----+-----+
## | School talent | Intelligence is important to get good   | 1\. Strongly disagree\  | 367 ( 6)  |
## |               | grades                                 | 2\. Disagree\          | 920 (15)  |
## |               |                                         | 3\. Agree\             | 2970 (48) |
## |               |                                         | 4\. Strongly agree     | 1760 (28) |
## +-----+-----+-----+-----+
## | School effort | Effort is important to get good grades | 1\. Strongly disagree\  | 109 ( 1)  |
## |               |                                         | 2\. Disagree\          | 88 ( 1)  |
## |               |                                         | 3\. Agree\             | 1427 (23) |
## |               |                                         | 4\. Strongly agree     | 4406 (73) |
## +-----+-----+-----+-----+
## | Social talent | In Chile, people are rewarded for their | 1\. Strongly disagree\  | 517 ( 9)  |
## |               | intelligence and skill                 | 2\. Disagree\          | 1568 (25) |
## |               |                                         | 3\. Agree\             | 2673 (43) |
## |               |                                         | 4\. Strongly agree     | 983 (17)  |
## +-----+-----+-----+-----+
## | Social effort | In Chile, people are rewarded for their | 1\. Strongly disagree\  | 512 ( 8)  |
## |               | efforts                               | 2\. Disagree\          | 1733 (28) |
## |               |                                         | 3\. Agree\             | 2607 (42) |
## |               |                                         | 4\. Strongly agree     | 1050 (17) |
## +-----+-----+-----+-----+
## | Deservingness | In Chile, people get what they deserve | 1\. Strongly disagree\  | 604 (10)  |
## |               |                                         | 2\. Disagree\          | 1911 (31) |
## |               |                                         | 3\. Agree\             | 2388 (39) |
## |               |                                         | 4\. Strongly agree     | 871 (15)  |
## +-----+-----+-----+-----+
## | Education      | Parental educational level             | 1\. 8th grade or less\  | 559 ( 8)  |
## |               |                                         | 2\. Secondary Education\ | 1698 (27) |
## |               |                                         | 3\. Higher tec. education\ | 960 (15)  |
## |               |                                         | 4\. University or Postgraduat\ | 1080 (17) |
## |               |                                         | 5\. Missing             | 1975 (32) |
## +-----+-----+-----+-----+
## | Books          | Number of books at home                | 1\. Les than 25\       | 3920 (63) |
## |               |                                         | 2\. More than 25       | 2281 (37) |
## +-----+-----+-----+-----+
## | Technology     | Technology access index                 | Mean (sd) : 7.8 (2.5)\  | 13 dist.  |
## |               |                                         | min < med < max:\     |           |

```



```

school$cod_depe2 <- set_label(school$cod_depe2, label = "Administrative dependency of school")

school$cod_grupo <- factor(school$cod_grupo, levels = c("Bajo", "Medio bajo", "Medio", "Medio alto", "Medio muy alto"))
school$cod_grupo <- set_label(school$cod_grupo, label = "Socioeconomic level of school")

df3<- dfSummary(school,
  plain.ascii = FALSE,
  style = "grid",
  tmp.img.dir = "/tmp",
  graph.magnif = 0.75,
  headings = F, # encabezado
  varnumbers = F, # num variable
  labels.col = T, # etiquetas
  na.col = F, # missing
  graph.col = F, # plot
  valid.col = T, # n valido
  col.widths = c(10,30,10,10,10))
df3$Variable <- c("Prop. university level
at school",
  "Achievement", "Dependency", "Socioeconomic level")
view(df3, footnote=NA, file = "output/tables/desc03.html")

## Output file written: C:\Users\kevin\OneDrive\Documentos\meritocracia-escuela-agencia\output\tables\d
webshot(url ="output/tables/desc03.html" ,file ="output/tables/desc03.png")

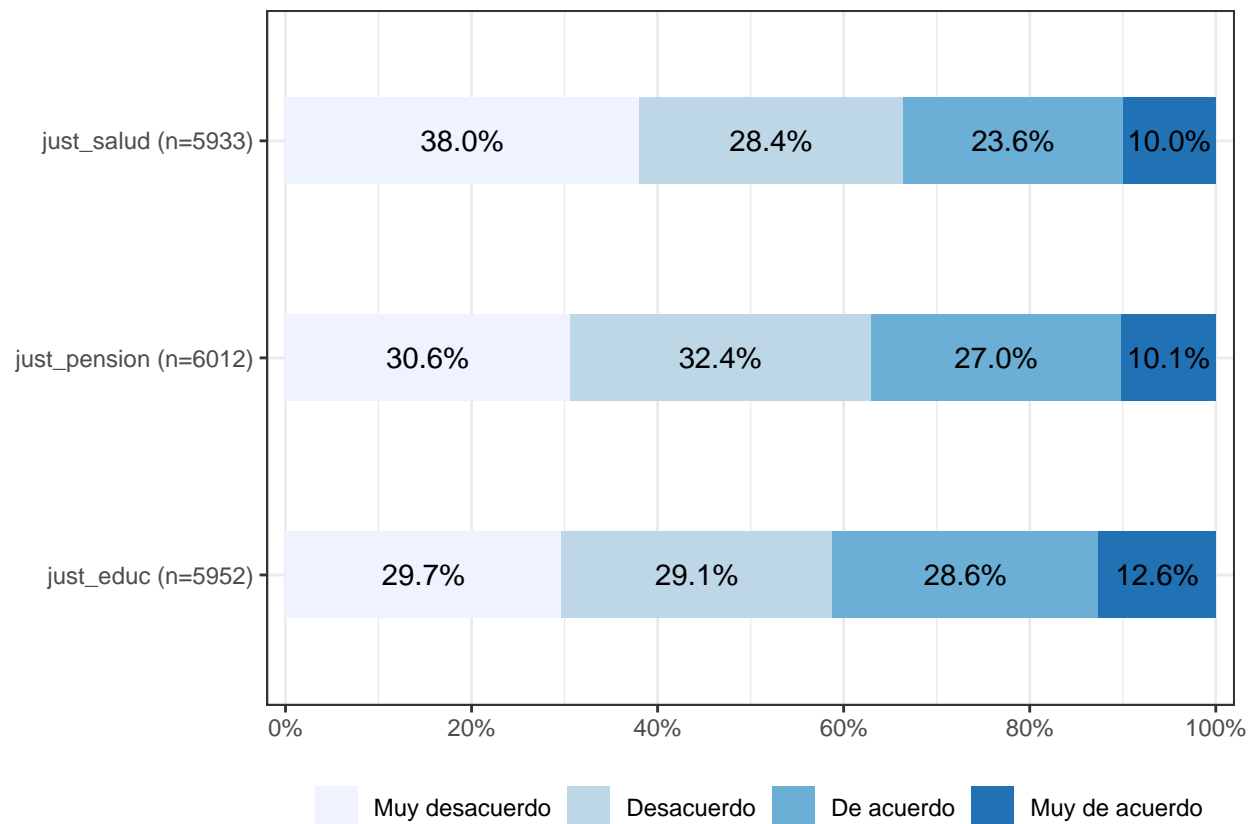
```

Variable	Label	Stats / Values	Freqs (% of Valid)	Valid
Prop. university level at school	Proportion of parents with university level by school	Mean (sd) : 0.2 (0.2) min ≤ med ≤ max: 0 ≤ 0.1 ≤ 0.9 IQR (CV) : 0.2 (0.9)	103 distinct values	6272 (100.0%)
Achievement	SIMCE score by school	1. Low 2. Medium 3. High	2091 (33.3%) 2091 (33.3%) 2090 (33.3%)	6272 (100.0%)
Dependency	Administrative dependency of school	1. Public 2. Subsidized private 3. Private	2659 (42.4%) 3169 (50.5%) 444 (7.1%)	6272 (100.0%)
Socioeconomic level	Socioeconomic level of school	1. Low 2. Medium low 3. Medium 4. Medium high 5. High	720 (11.5%) 2282 (36.4%) 1383 (22.1%) 1309 (20.9%) 578 (9.2%)	6272 (100.0%)

graficos descriptivos

dependientes

```
dependientes <- sjPlot::plot_stackfrq(dplyr::select(data, just_educ, just_pension, just_salud), geom.si
  theme_bw() +
  theme(legend.position="bottom")
dependientes
```



```
ggsave(dependientes, file = "output/graphs/dependientes.png",device = "png",width = 25,height = 13,dpi = 300)
```

independientes

```
independientes <- sjPlot::plot_stackfrq(dplyr::select(data, inteligencia_esc, esfuerzo_esc, inteligencia_esc),
  theme_bw() +
  theme(legend.position="bottom")
)
ggsave(independientes, file = "output/graphs/independientes.png",device = "png",width = 25,height = 13,dpi = 300)
```