

indice

Eduardo Bologna

28/12/2020

Lectura de las bases de datos

```
eph_1_05<-read.dbf("bases_EPH/Ind_t105.DBF")
eph_1_10<-read.dbf("bases_EPH/Ind_t110.dbf")
eph_1_15<-read.dbf("bases_EPH/Ind_t115.dbf")
eph_1_20<-read.table("bases_EPH/usu_individual_T120.txt",
                     header = TRUE, sep= ";", dec = ",")
```

Selección de la base a usar

```
base<-eph_1_20
```

Organización

rotular variables, crear grupos de edades, recodificar y eliminar valores perdidos

```
base$REGION<-as.factor(base$REGION)
levels(base$REGION)<-c("Gran Buenos Aires", "NOA",
                      "NEA", "Cuyo", "Pampeana", "Patagónica")

base$CH04<-as.factor(base$CH04)
levels(base$CH04)<-c("varones", "mujeres")

base$CH06[base$CH06==1]<-0
base$grupos_edades<-cut(base$CH06, c(0,24,39, 54,99))

base$NIVEL_ED[base$NIVEL_ED==9]<-NA
base$NIVEL_ED[base$NIVEL_ED==7]<-0
base$NIVEL_ED<-as.factor(base$NIVEL_ED)
levels(base$NIVEL_ED)<-c("primario incompleto o menos", "primario incompleto o menos", "primario completo",
                        "secundario incompleto", "secundario completo",
                        "superior incompleto", "superior completo")
```

Tasa de actividad

solo mayores de 9 años

```
mayores_9<-subset(base, base$CH06>9)
mayores_9$ESTADO[mayores_9$ESTADO==0]<-NA

tasa_actividad <- (((addmargins(wtd.table(mayores_9$ESTADO, weights = mayores_9$PONDERA))) [1]+
                    (addmargins(wtd.table(mayores_9$ESTADO, weights = mayores_9$PONDERA))) [2])/
                    (addmargins(wtd.table(mayores_9$ESTADO, weights = mayores_9$PONDERA))) [4]))
tasa_actividad<-as.numeric(tasa_actividad)
```

Por grupos

```
actividad_sexo<-mayores_9 %>%
  group_by(CH04) %>%
  summarise(valor=(((addmargins(wtd.table(ESTADO, weights = PONDERA))) [1]+
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [2])/
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [4])))

actividad_region<-mayores_9 %>%
  group_by(REGION) %>%
  summarise(valor=(((addmargins(wtd.table(ESTADO, weights = PONDERA))) [1]+
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [2])/
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [4])))

actividad_edad<-mayores_9 %>%
  group_by(grupos_edades) %>%
  summarise(valor=(((addmargins(wtd.table(ESTADO, weights = PONDERA))) [1]+
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [2])/
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [4])))

actividad_educacion<-mayores_9 %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=(((addmargins(wtd.table(ESTADO, weights = PONDERA))) [1]+
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [2])/
                    (addmargins(wtd.table(ESTADO, weights = PONDERA))) [4])))
```

Gaps

```
gap_actividad_sexo<-100*round((actividad_sexo[1,2]-
                              actividad_sexo[2,2])/actividad_sexo[1,2],4)

gap_actividad_region<-100*round((max(actividad_region[,2])-
                                   min(actividad_region[,2]))/max(actividad_region[,2]),4)

gap_actividad_edad<-100*round((max(actividad_edad[,2])-
                                   min(actividad_edad[,2]))/max(actividad_edad[,2]),4)
```

```
gap_actividad_educacion<-100*round((max(actividad_educacion[,2])-
min(actividad_educacion[,2]))/max(actividad_educacion[,2]),4)
```

Desocupación

solo personas económicamente activas

```
activos<-subset(base, base$ESTADO==1 | base$ESTADO==2)

tasa_desocupacion <- addmargins(wtd.table(activos$ESTADO, weights = activos$PONDERA))[2] /
addmargins(wtd.table(activos$ESTADO, weights = activos$PONDERA))[3]

tasa_desocupacion<-as.numeric(tasa_desocupacion)
```

Por grupos

```
desocupacion_sexo<-activos %>%
  group_by(CH04) %>%
  summarise(valor=(addmargins(wtd.table(ESTADO, weights = PONDERA))[2] /
addmargins(wtd.table(ESTADO, weights = PONDERA))[3]))

desocupacion_region<-activos %>%
  group_by(REGION) %>%
  summarise(valor=(addmargins(wtd.table(ESTADO, weights = PONDERA))[2] /
addmargins(wtd.table(ESTADO, weights = PONDERA))[3]))

desocupacion_edad<-activos %>%
  group_by(grupos_edades) %>%
  summarise(valor=(addmargins(wtd.table(ESTADO, weights = PONDERA))[2] /
addmargins(wtd.table(ESTADO, weights = PONDERA))[3]))

desocupacion_educacion<-activos %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=(addmargins(wtd.table(ESTADO, weights = PONDERA))[2] /
addmargins(wtd.table(ESTADO, weights = PONDERA))[3]))
```

Gaps

```
gap_desocupacion_sexo<-100*round((desocupacion_sexo[1,2]-desocupacion_sexo[2,2])/
desocupacion_sexo[1,2],4)

gap_desocupacion_region<-100*round((max(desocupacion_region[,2])-min(desocupacion_region[,2]))/
max(desocupacion_region[,2]),4)

gap_desocupacion_edad<-100*round((max(desocupacion_edad[,2])-min(desocupacion_edad[,2]))/
```

```

max(desocupacion_edad[,2]),4)

gap_desocupacion_educacion<-100*round((max(desocupacion_educacion[,2])-min(desocupacion_educacion[,2]))/
max(desocupacion_educacion[,2]),4)

```

Proporción de personas que trabajan por cuenta propia

solo personas ocupadas

```

ocupades<-subset(base, base$ESTADO==1)
prop_independientes<-(addmargins(wtd.table(ocupades$CAT_OCUP, weights = ocupades$PONDERA))) [2] /
  (addmargins(wtd.table(ocupades$CAT_OCUP, weights = ocupades$PONDERA))) [5]

prop_independientes<-as.numeric(prop_independientes)

```

Por grupos

```

prop_independientes_sexo<-ocupades %>%
  group_by(CH04) %>%
  summarise(valor=(addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [2] /
    (addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [5])

prop_independientes_region<-ocupades %>%
  group_by(REGION) %>%
  summarise(valor=(addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [2] /
    addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [5])

prop_independientes_edad<-ocupades %>%
  group_by(grupos_edades) %>%
  summarise(valor=(addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [2] /
    addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [5])

prop_independientes_educacion<-ocupades %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=(addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [2] /
    addmargins(wtd.table(CAT_OCUP, weights = PONDERA))) [5])

```

Gaps

```

gap_prop_independientes_sexo<-100*round(
  (prop_independientes_sexo[1,2]-
    prop_independientes_sexo[2,2])/prop_independientes_sexo[1,2],4)

gap_prop_independientes_region<-100*round(
  (max(prop_independientes_region[,2])-
    min(prop_independientes_region[,2]))/max(prop_independientes_region[,2]),4)

```

```

gap_prop_independientes_edad<-100*round(
  (max(prop_independientes_edad[,2])-
    min(prop_independientes_edad[,2]))/max(prop_independientes_edad[,2]),4)

gap_prop_independientes_educacion<-100*round(
  (max(prop_independientes_educacion[,2])-
    min(prop_independientes_educacion[,2]))/max(prop_independientes_educacion[,2]),4)

```

Seguridad

solo personas asalariadas

estabilidad + obra social +aguinaldo + vacaciones + dias por enfermedad + descuento jubilatorio

```

asalariades<-subset(base, base$CAT_OCUP==3)

asalariades$PP07C[asalariades$PP07C==0]<-NA
asalariades$PP07C[asalariades$PP07C==9]<-NA
asalariades$estabilidad<-asalariades$PP07C-1

asalariades$PP07G1[asalariades$PP07G1==0]<-NA
asalariades$PP07G1[asalariades$PP07G1==9]<-NA
asalariades$vacaciones <-asalariades$PP07G1
asalariades$vacaciones[asalariades$vacaciones==2]<-0

asalariades$PP07G2[asalariades$PP07G2==0]<-NA
asalariades$PP07G2[asalariades$PP07G2==9]<-NA
asalariades$aguinaldo<-asalariades$PP07G2
asalariades$aguinaldo[asalariades$aguinaldo==2]<-0

asalariades$PP07G3[asalariades$PP07G3==0]<-NA
asalariades$PP07G3[asalariades$PP07G3==9]<-NA
asalariades$dias_enfermedad<-asalariades$PP07G3
asalariades$dias_enfermedad[asalariades$dias_enfermedad==2]<-0

asalariades$PP07G4[asalariades$PP07G4==0]<-NA
asalariades$PP07G4[asalariades$PP07G4==9]<-NA
asalariades$obra_social<-asalariades$PP07G4
asalariades$obra_social[asalariades$obra_social==2]<-0

asalariades$PP07H[asalariades$PP07H==0]<-NA
asalariades$PP07H[asalariades$PP07H==9]<-NA
asalariades$aporte_jubilatorio<-asalariades$PP07H
asalariades$aporte_jubilatorio[asalariades$aporte_jubilatorio==2]<-0

asalariades$seguridad<-(asalariades$estabilidad + asalariades$vacaciones+
  asalariades$aguinaldo + asalariades$dias_enfermedad+

```

```
asalariades$sobra_social+
asalariades$aporte_jubilatorio)/6
```

```
seguridad<-round(wtd.mean(asalariades$seguridad, weights = asalariades$PONDERA),3)
sd_seguridad<-weightedSd(asalariades$seguridad, weights = asalariades$PONDERA,
na.rm=TRUE)
```

Por grupos

```
seguridad_sexo<-asalariades %>%
  group_by(CH04) %>%
  summarise(valor=wtd.mean(seguridad, weights = PONDERA))

seguridad_region<-asalariades %>%
  group_by(REGION) %>%
  summarise(valor=wtd.mean(seguridad, weights = PONDERA))

seguridad_edad<-asalariades %>%
  group_by(grupos_edades) %>%
  summarise(valor=wtd.mean(seguridad, weights = PONDERA))

seguridad_educacion<-asalariades %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=wtd.mean(seguridad, weights = PONDERA))
```

Gaps

```
gap_seguridad_sexo<-100*round(
  (seguridad_sexo[1,2]-seguridad_sexo[2,2])/
  seguridad_sexo[1,2],4)

gap_seguridad_region<-100*round(
  (max(seguridad_region[,2])-min(seguridad_region[,2]))/
  max(seguridad_region[,2]),4)

gap_seguridad_edad<-100*round(
  (max(seguridad_edad[,2])-min(seguridad_edad[,2]))/
  max(seguridad_edad[,2]),4)

gap_seguridad_educacion<-100*round(
  (max(seguridad_educacion[,2])-min(seguridad_educacion[,2]))/
  max(seguridad_educacion[,2]),4)
```

Ingresos salariales

solo personas ocupadas con ingreso no nulo

```
ocupades$PP08D1[ocupades$PP08D1==9]<-NA
ocupades$PP08D1[ocupades$PP08D1==0]<-NA
ocupades$ingresos<-ocupades$PP08D1
ingresos<-wtd.mean(ocupades$ingresos, weights = ocupades$PONDERA)
```

Por grupos

```
ingreso_sexo<-ocupades %>%
  group_by(CH04) %>%
  summarise(valor=wtd.mean(ingresos, weights = PONDERA))

ingreso_region<-ocupades %>%
  group_by(REGION) %>%
  summarise(valor=wtd.mean(ingresos, weights = PONDERA))

ingreso_edad<-ocupades %>%
  group_by(grupos_edades) %>%
  summarise(valor=wtd.mean(ingresos, weights = PONDERA))

ingreso_educacion<-ocupades %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=wtd.mean(ingresos, weights = PONDERA))
```

Gaps

```
gap_ingreso_sexo<-100*round((ingreso_sexo[1,2]-ingreso_sexo[2,2])/
                             ingreso_sexo[1,2],4)

gap_ingreso_region<-100*round((max(ingreso_region[,2])-min(ingreso_region[,2]))/
                               max(ingreso_region[,2]),4)

gap_ingreso_edad<-100*round((max(ingreso_edad[,2])-min(ingreso_edad[,2]))/
                             max(ingreso_edad[,2]),4)

gap_ingreso_educacion<-100*round((max(ingreso_educacion[,2])-min(ingreso_educacion[,2]))/
                                  max(ingreso_educacion[,2]),4)
```

Ingreso - hora

```
ocupades$horas_semana<-ocupades$PP3E_TOT
ocupades$horas_semana[ocupades$horas_semana==0]<-NA
```

```
ocupades$horas_semana[ocupades$horas_semana>84]<-NA

ocupades$ingreso_hora<-ocupades$ingresos/(ocupades$horas_semana*4)

ingreso_hora<-wtd.mean(ocupades$ingreso_hora, weights = ocupades$PONDERA)
```

Por grupos

```
ingreso_hora_sexo<-ocupades %>%
  group_by(CH04) %>%
  summarise(valor=wtd.mean(ingreso_hora, weights = PONDERA))

ingreso_hora_region<-ocupades %>%
  group_by(REGION) %>%
  summarise(valor=wtd.mean(ingreso_hora, weights = PONDERA))

ingreso_hora_edad<-ocupades %>%
  group_by(grupos_edades) %>%
  summarise(valor=wtd.mean(ingreso_hora, weights = PONDERA))

ingreso_hora_educacion<-ocupades %>%
  group_by(NIVEL_ED) %>%
  summarise(valor=wtd.mean(ingreso_hora, weights = PONDERA))
```

Gaps

```
gap_ingreso_hora_sexo<-100*round((ingreso_hora_sexo[1,2]-ingreso_hora_sexo[2,2])/
                                ingreso_hora_sexo[1,2],4)

gap_ingreso_hora_region<-100*round((max(ingreso_hora_region[,2])-min(ingreso_hora_region[,2]))/
                                max(ingreso_hora_region[,2]),4)

gap_ingreso_hora_edad<-100*round((max(ingreso_hora_edad[,2])-min(ingreso_hora_edad[,2]))/
                                max(ingreso_hora_edad[,2]),4)

gap_ingreso_hora_educacion<-100*round((max(ingreso_hora_educacion[,2])-min(ingreso_hora_educacion[,2]))/
                                max(ingreso_hora_educacion[,2]),4)
```

Cociente de la renta media per cápita del 10% más rico al 50% más pobre

```
base$IPCF[base$IPCF==0]<-NA

ingresos_bajos<-subset(
  base, base$IPCF<=quantile(base$IPCF,.50, na.rm = TRUE))
ingresos_altos<-subset(
```



```

base, base$IPCF>=quantile(base$IPCF,.90, na.rm = TRUE))
p90_p50_total<-wtd.mean(
  ingresos_altos$IPCF, weights = ingresos_altos$PONDERA)/
  wtd.mean(ingresos_bajos$IPCF, weights = ingresos_bajos$PONDERA)

```

Por grupos

```

ingresos_bajos_sexo<-base %>% group_by(CH04) %>%
  subset(IPCF<=quantile(IPCF,.50, na.rm = TRUE))%>%
  summarise(media_ingresos_bajos=wtd.mean(IPCF, weights = PONDERA))

ingresos_altos_sexo<-base %>% group_by(CH04) %>%
  subset(IPCF>=quantile(IPCF,.90, na.rm = TRUE))%>%
  summarise(media_ingresos_altos=wtd.mean(IPCF, weights = PONDERA))

p90_p50_sexos<-merge(ingresos_bajos_sexo, ingresos_altos_sexo)
p90_p50_sexos$diferencia<-p90_p50_sexos$media_ingresos_altos/p90_p50_sexos$media_ingresos_bajos

ingresos_bajos_region<-base %>% group_by(REGION) %>%
  subset(IPCF<=quantile(IPCF,.50, na.rm = TRUE))%>%
  summarise(media_ingresos_bajos=wtd.mean(IPCF, weights = PONDERA))

ingresos_altos_region<-base %>% group_by(REGION) %>%
  subset(IPCF>=quantile(IPCF,.90, na.rm = TRUE))%>%
  summarise(media_ingresos_altos=wtd.mean(IPCF, weights = PONDERA))

p90_p50_region<-merge(ingresos_bajos_region, ingresos_altos_region)
p90_p50_region$diferencia<-p90_p50_region$media_ingresos_altos/p90_p50_region$media_ingresos_bajos

ingresos_bajos_edad<-base %>% group_by(grupos_edades) %>%
  subset(IPCF<=quantile(IPCF,.50, na.rm = TRUE))%>%
  summarise(media_ingresos_bajos=wtd.mean(IPCF, weights = PONDERA))

ingresos_altos_edad<-base %>% group_by(grupos_edades) %>%
  subset(IPCF>=quantile(IPCF,.90, na.rm = TRUE))%>%
  summarise(media_ingresos_altos=wtd.mean(IPCF, weights = PONDERA))

p90_p50_edad<-merge(ingresos_bajos_edad, ingresos_altos_edad)
p90_p50_edad$diferencia<-p90_p50_edad$media_ingresos_altos/p90_p50_edad$media_ingresos_bajos

ingresos_bajos_educacion<-base %>% group_by(NIVEL_ED) %>%
  subset(IPCF<=quantile(IPCF,.50, na.rm = TRUE))%>%
  summarise(media_ingresos_bajos=wtd.mean(IPCF, weights = PONDERA))

ingresos_altos_educacion<-base %>% group_by(NIVEL_ED) %>%
  subset(IPCF>=quantile(IPCF,.90, na.rm = TRUE))%>%
  summarise(media_ingresos_altos=wtd.mean(IPCF, weights = PONDERA))

```

```
p90_p50_educacion<-merge(ingresos_bajos_educacion, ingresos_altos_educacion)
p90_p50_educacion$diferencia<-p90_p50_educacion$media_ingresos_altos/p90_p50_educacion$media_ingresos_bajos
```

Comparaciones simples

```
sexo<-c("varones", "mujeres")
region<-c("Gran Buenos Aires", "NOA",
          "NEA", "Cuyo", "Pampeana", "Patagónica")
edad<-c("menos de 25", "25 -39", "40 - 54", "55 y más")
educacion<-c("primario incompleto o menos", "primario completo",
             "secundario incompleto", "secundario completo",
             "superior incompleto", "superior completo")

grupos<-c("general", sexo, region, edad, educacion)

valores_actividad<-c(tasa_actividad[[1]], actividad_sexo[[2]],
                     actividad_region[[2]], actividad_edad[[2]],
                     actividad_educacion[[2]])

valores_desocupacion<-c(tasa_desocupacion[[1]], desocupacion_sexo[[2]],
                        desocupacion_region[[2]], desocupacion_edad[[2]],
                        desocupacion_educacion[[2]])

valores_prop_independientes<- c(prop_independientes[[1]], prop_independientes_sexo[[2]],
                                prop_independientes_region[[2]], prop_independientes_edad[[2]],
                                prop_independientes_educacion[[2]])

valores_seguridad<-c(seguridad[[1]], seguridad_sexo[[2]], seguridad_region[[2]],
                     seguridad_edad[[2]], seguridad_educacion[[2]])

valores_ingreso<-c(ingresos[[1]], ingreso_sexo[[2]], ingreso_region[[2]],
                   ingreso_edad[[2]], ingreso_educacion[[2]])

valores_ingreso_hora<-c(ingreso_hora[[1]], ingreso_hora_sexo[[2]], ingreso_hora_region[[2]],
                        ingreso_hora_edad[[2]], ingreso_hora_educacion[[2]])

valores_diferencia_p90_p50<-c(p90_p50_sexos$diferencia, p90_p50_region$diferencia,
                              p90_p50_edad$diferencia, p90_p50_educacion$diferencia)

compara_grupos<-data.frame(grupos, 100*round(valores_actividad,3),
                           100*round(valores_desocupacion,3),
                           100*round(valores_prop_independientes,3),
                           round(valores_seguridad,3), round(valores_ingreso,2),
                           round(valores_ingreso_hora,2),
                           round(valores_diferencia_p90_p50,1))
names(compara_grupos)<-c("grupo", "tasa de actividad", "tasa de desocupación",
                        "proporción de trabajadores por cuenta propia", "seguridad",
                        "ingreso medio", "ingreso-hora medio",
                        "renta media 10% más rico a 50% más pobre")

kable(compara_grupos, format="latex", booktabs=TRUE) %>%
```

```
kable_styling(latex_options="scale_down")%>%
column_spec(2:17, width = "2cm")%>%
  row_spec(0, align = "c") # qué paso que repite columnas???
```

[illegible]

Normalización de los valores al intervalo [0 - 100]

[illegible]

Labor Unequality Index (LaUnIn)

Índice aditivo

```
compara_grupos_normalized$LaUnIn<-round((compara_grupos_normalized$tasa.de.actividad.1+  
      (100-compara_grupos_normalized$tasa.de.desocupación.1)+  
      (100-compara_grupos_normalized$proporción.de.trabajadores.por.cuenta.propia.1)+  
        compara_grupos_normalized$seguridad.1+  
        compara_grupos_normalized$ingreso.medio.1+  
        compara_grupos_normalized$ingreso.hora.medio.1+ (100-compara_grupos_normalized
```

Gráfico

```
compara_grupos_normalizado$fronteras<-c("media nacional", rep("sexo",2), rep("region", 6), rep("edad",4))
compara_grupos_normalizado$grupo <- factor(compara_grupos_normalizado$grupo,
                                             levels = compara_grupos_normalizado$grupo)
ggplot(compara_grupos_normalizado[-1,] )+
  geom_bar(aes(grupo, LaUnIn, fill=fronteras), stat = "identity")+
  geom_hline(yintercept = compara_grupos_normalizado[1,16], col="red")+coord_flip()+
  theme_tufte()
```

