

Argentina: brecha de género

Eduardo Bologna

10/3/2020

```
library(foreign)
```

```
## Warning: package 'foreign' was built under R version 3.5.3
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
library(ggthemes)  
library(doBy)  
library(scales)
```

```
## Warning: package 'scales' was built under R version 3.5.3
```

Se compara la evolución de la inserción laboral de mujeres y varones en los aglomerados urbanos de Argentina durante el período 2004-2018. Para homogeneizar la comparación, solo se toma el primer semestre de cada año.

Preliminares

El INDEC pone a disposición las bases de microdatos de la EPH en su sitio <https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos>. Entre 2004 y 2015 las bases están en formato .dbf y los tres últimos años del período en .txt, por lo que, para la lectura de los archivos, se definen dos listas que contienen las ubicaciones de las bases en cada uno de los formatos.

```
ubicaciones.bases.dbf=list(  
  "archivosINDEC\\Ind_t104.dbf", "archivosINDEC\\Ind_t105.dbf",  
  "archivosINDEC\\Ind_t106.dbf", "archivosINDEC\\Ind_t107.dbf",  
  "archivosINDEC\\Ind_t108.dbf", "archivosINDEC\\Ind_t109.dbf",  
  "archivosINDEC\\Ind_t110.dbf", "archivosINDEC\\Ind_t111.dbf",  
  "archivosINDEC\\Ind_t112.dbf", "archivosINDEC\\Ind_t113.dbf",  
  "archivosINDEC\\Ind_t114.dbf", "archivosINDEC\\Ind_t115.dbf")  
  
ubicaciones.bases.txt=list(  
  "archivosINDEC\\usu_individual_t316.txt",  
  "archivosINDEC\\usu_individual_t117.txt",  
  "archivosINDEC\\usu_individual_t118.txt")
```

Y dos vectores que contienen las matrices de datos:

```
eph.dbf=vector("list",12)
for(i in 1:12) eph.dbf[[i]]=read.dbf(ubicaciones.bases.dbf[[i]])

eph.txt=vector("list",3)
for(i in 1:3) eph.txt[[i]]=read.csv(ubicaciones.bases.txt[[i]], sep = ";")
```

En los últimos tres años los ingresos salariales nulos o perdidos fueron dejados en blanco, por lo que se los convierte en ceros

```
for(i in 1:3) eph.txt[[i]]$PP08D1[is.na(eph.txt[[i]]$PP08D1)]=0
```

Se reúnen todas las bases en un único vector

```
eph=vector("list",15)
eph=c(eph.dbf,eph.txt)
```

Se define `sexo` como factor y sus niveles, se restringen las bases a los casos válidos de la variable `ESTADO` y a los mayores ed 14 años de edad, se define `actividad` como la rotulación de `ESTADO`, se define `educacion` como `NIVEL_ED` con el orden corregido y rotulada.

```
for (i in 1:15) {
  eph[[i]]$sexo=as.factor(eph[[i]]$CH04)
  levels(eph[[i]]$sexo)=c("varones", "mujeres")
  eph[[i]]=subset(eph[[i]], eph[[i]]$ESTADO!=0 & eph[[i]]$ESTADO!=4 &
    eph[[i]]$CH06>14)
  eph[[i]]$actividad=as.factor(eph[[i]]$ESTADO)
  levels(eph[[i]]$actividad)=c("ocupado", "desocupado", "inactivo")
  eph[[i]]$educacion=factor(eph[[i]]$NIVEL_ED,
    levels(factor(eph[[i]]$NIVEL_ED))[c(7,1:6)])
  levels(eph[[i]]$educacion)=c("Sin instrucción",
    "Primaria Incompleta",
    "Primaria Completa","Secundaria Incompleta", "Secundaria Completa",
    "Universitaria Incompleta","Universitaria Completa")
}
```

Nivel de actividad económica

Se define `conduct` como 1 para activos y 0 para inactivos a partir de `ESTADO`

```
for (i in 1:15) {
  eph[[i]]$conduct=ifelse(eph[[i]]$ESTADO==3, 0, 1)
}

for (i in 1:15) {
  eph[[i]]$conduct=factor(eph[[i]]$conduct)
  levels(eph[[i]]$conduct)=c("inactivos", "activos")
}
```

Se define una nueva matriz de datos a partir de la tabla que cruza condición de actividad, educación y sexo y se le agrega la proporción de activos en cada categoría de sexo y educación:

```
prop.activos=vector("list",15)
for (i in 1:15) {
  prop.activos[[i]]=as.data.frame(ftable(prop.table(
    table(eph[[i]]$condact, eph[[i]]$educacion, eph[[i]]$sexo),
    c(3,2))))
  names(prop.activos[[i]])=c("condicion", "educacion","sexo", "proporcion")
}
```

Se define el vector que alojará los gráficos de esta matriz y se construyen gráficos de barras

```
graficos.condact=vector("list",15)

for (i in 1:15) {
  graficos.condact[[i]]=ggplot(eph[[i]])+
    geom_bar(aes(x=educacion, fill=factor(condact,
      position = "fill")+
    scale_fill_brewer(type = "div",
      palette = "Set2", direction = 1,
      aesthetics = "fill")+
    guides(fill=guide_legend(title=NULL))+
    geom_text(
      data=prop.activos[[i]],position=position_stack(vjust=.5),
      size=3, aes(x= educacion,
        y=proporcion,label =paste(100*round(proporcion,3),"%")))+
    labs(
      title="Condición de actividad según nivel de educación y sexo",
      subtitle = paste("Argentina. Aglomerados Urbanos primer trimestre",eph[[i]]$AN04),
      caption="Fuente: Encuesta Permanente de Hogares, INDEC www.indec.gob.ar",
      x="Máximo nivel de educación alcanzado"
    )+theme_economist(base_family = "serif", base_size = 9)+
    theme(axis.line=element_blank(),axis.text.y=element_blank(),
      axis.title.y=element_blank(),
      axis.ticks=element_blank(),legend.title=element_blank(),
      plot.title = element_text(hjust = 0.5))+
    facet_grid(sexo~.)
}
```

levels=c("a

Y se guardan como imágenes en una carpeta exclusiva

```
for(i in 1:15){
  ggsave(filename = paste("condiciondeactividad", eph[[i]]$AN04,
    ".png"),
    path = "graficos\\condicion_actividad",
    plot = graficos.condact[[i]], device = "png",
    scale = 1.7)
}
```

```
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
```

```
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
```

Desocupación

Se restringe la matriz de datos al grupo de quienes tienen actividad económica

```
solo.activos=vector("list", 15)
for (i in 1:15) {
  solo.activos[[i]]=subset(eph[[i]],
                           eph[[i]]$ESTADO==1 | eph[[i]]$ESTADO==2)
  solo.activos[[i]]$actividad=factor(solo.activos[[i]]$actividad)
}
```

Se generan matrices de datos con la proporción de desocupados respecto del total de activos, según sexo y educación.

```
props.d=vector("list",15)
for (i in 1:15) {
  props.d[[i]]=as.data.frame(ftable(prop.table(
    table(solo.activos[[i]]$actividad, solo.activos[[i]]$educacion,
      solo.activos[[i]]$sexo),
    c(3,2))))
  names(props.d[[i]])=
    c("ocupacion", "educacion", "sexo", "proporcion")
}
```

Se construyen los gráficos de barras y se guardan en su propia carpeta

```
graficos.desocupacion=vector("list", 15)
for (i in 1:15) {
  graficos.desocupacion[[i]]=ggplot(solo.activos[[i]])+
    geom_bar(aes(x=educacion, fill=factor(actividad,
                                          levels=c("desocupado",
                                                    "ocupado"))),
             position = "fill")+
    scale_fill_brewer(type = "div", palette = "Set2", direction = 1,
                      aesthetics = "fill")+
    guides(fill=guide_legend(title=NULL))+
    geom_text(data=props.d[[i]],
              position=position_stack(vjust=0.5), size=3,
              aes(x= educacion,y=proporcion,
                  label =paste(100*round(proporcion,3),"%")))+
  labs(
```

```

    title="Condición de ocupación de la población económicamente activa
    según nivel de educación y sexo",
    subtitle = paste("Argentina. Aglomerados Urbanos primer trimestre",eph[[i]]$ANO4),
    caption="Fuente: Encuesta Permanente de Hogares, INDEC www.indec.gob.ar",
    x="Máximo nivel de educación alcanzado"
)+theme_economist(base_family = "serif", base_size = 9)+
theme(axis.line=element_blank(),axis.text.y=element_blank(),
      axis.title.y=element_blank(),
      axis.ticks=element_blank(),legend.title=element_blank(),
      plot.title = element_text(hjust = 0.5))+
facet_grid(sexo~.)
}

for(i in 1:15){
  ggsave(filename = paste("desocupación", eph[[i]]$ANO4,
                          ".png"),
        path = "graficos\\desocupacion",
        plot = graficos.desocupacion[[i]], device = "png",
        scale = 1.7)
}

```

```

## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image

```

Categoría ocupacional

De la base que solo contiene a las personas económicamente activas, se retienen quienes declaran alguna de las tres categorías ocupacionales y se rotulan

```

solo.activos.3.cats=vector("list", 15)
for (i in 1:15) {
  solo.activos.3.cats[[i]]=subset(solo.activos[[i]], solo.activos[[i]]$CAT_OCUP==1 |
                                solo.activos[[i]]$CAT_OCUP==2 |
                                solo.activos[[i]]$CAT_OCUP==3)
  solo.activos.3.cats[[i]]$categoria=
    as.factor(as.character(solo.activos.3.cats[[i]]$CAT_OCUP))
  levels(solo.activos.3.cats[[i]]$categoria)=c("Patrón", "Cuenta propia","Obrero o empleado")
}

```

Se construyen las bases con las proporciones por sexo y educación, se grafica y se guarda

```

props.cats=vector("list",15)
for (i in 1:15) {
  props.cats[[i]]=as.data.frame(ftable(prop.table(
    table(solo.activos.3.cats[[i]]$categoria,
          solo.activos.3.cats[[i]]$educacion,
          solo.activos.3.cats[[i]]$sexo),
    c(3,2))))
  names(props.cats[[i]])=
    c("ocupacion", "educacion","sexo", "proporcion")
}

graficos.categoria=vector("list", 15)
for (i in 1:15) {
  graficos.categoria[[i]]=ggplot(solo.activos.3.cats[[i]])+
    geom_bar(aes(x=educacion, fill=factor(categoria,
                                          levels=c("Patrón", "Cuenta propia",
                                                    "Obrero o empleado"))),
            position = "fill")+
    scale_fill_brewer(type = "div", palette = "Set2", direction = 1,
                      aesthetics = "fill")+
    guides(fill=guide_legend(title=NULL))+
    geom_text(data=props.cats[[i]],
              position=position_stack(vjust=0.5), size=3,
              aes(x= educacion,y=proporcion,group=ocupacion,
                  label =paste(100*round(proporcion,3),"%")))+
  labs(
    title="Categoría ocupacional según nivel de educación y sexo",
    subtitle = paste("Argentina. Aglomerados Urbanos primer trimestre",eph[[i]]$ANO4),
    caption="Fuente: Encuesta Permanente de Hogares, INDEC www.indec.gob.ar",
    x="Máximo nivel de educación alcanzado"
  )+theme_economist(base_family = "serif", base_size = 9)+
  theme(axis.line=element_blank(),axis.text.y=element_blank(),
        axis.title.y=element_blank(),
        axis.ticks=element_blank(),legend.title=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  facet_grid(sexo~.)
}

for(i in 1:15){
  ggsave(filename = paste("categoria", eph[[i]]$ANO4,
                          ".png"),
          path = "graficos\\categoria_ocupacional",
          plot = graficos.categoria[[i]], device = "png",
          scale = 1.7)
}

## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image

```

```
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
## Saving 11 x 7.65 in image
```