

# TALLER: UNA GUÍA PRÁCTICA PARA USAR LOS DATOS DEL BARÓMETRO DE LAS AMÉRICAS

Arturo Maldonado

22/09/2020

## REPLICACIÓN DEL TOPICAL BRIEF 035

### Tolerancia a los “golpes de Estado” ejecutivos en Perú

#### ABSTRACT

En este taller vamos a replicar el informe Actualidad #035 de la Serie Perspectivas desde el Barómetro de las Américas. Se puede ver aquí la lista de los reportes aquí. El informe 142 se puede conseguir en español aquí.

Para este taller, se asume que los participantes tienen un conocimiento básico de R y RStudio. A lo largo de este taller, explicaremos, paso a paso, los procedimientos en RStudio para producir las tablas y gráficos que se ven en los informes. En estos informes, estas tablas y gráficos han sido producidos originalmente en STATA.

En resumen, este informe presenta el promedio de tolerancia a los “golpes de Estado” ejecutivos en las Américas y la tendencia en el tiempo de esta tolerancia en Perú.

#### SOBRE LA BASE DE DATOS

Los datos que vamos a usar deben citarse de la siguiente manera: Fuente: Barómetro de las Américas por el Proyecto de Opinión Pública de América Latina (LAPOP), [www.LapopSurveys.org](http://www.LapopSurveys.org).

Pueden descargar los datos de manera libre aquí

En este enlace, se puede registrar o entrar como “Free User”. En el buscador, se puede ingresar el texto “merge”. Ahí se tendrá acceso a la base de datos completa “2004-2018 Grand Merge Free” en versión para STATA. Se descarga la base de datos en formato zip, la que se descomprime en formato .dta.

Debido a que estas bases de datos son muy grandes (42MB en formato zip), para este taller vamos a trabajar con una versión reducida, que incluye los países y años seleccionados en el informe. Esta base de datos y todos los materiales para esta taller se pueden descargar aquí

#### INSTALANDO LOS PAQUETES NECESARIOS EN R

#### LEYENDO LA BASE DE DATOS EN R

*Si se trabajara con la base de datos descargada del repositorio de LAPOP*

Se carga la base de datos en STATA en R en un dataframe “lapopmerge”. Tenga en cuenta que la base de datos tiene que estar en su directorio de trabajo (working directory). Para este Insights se analiza la ronda 2016/17, por lo que se selecciona esta ronda en el merge total. Se crea un nuevo dataframe “lapop”. Luego, también se elimina las observaciones de los países que no se analizan en este reporte por no tener datos de la variable sobre medios de comunicación. Se elimina el dataframe del Merge original. Finalmente el dataframe “lapop” incluye solo la ronda 2016/17 y los países de análisis.

Como en este taller estamos trabajando con la versión simplificada de la base de datos, directamente leemos esta base de datos en RStudio.

```
lapop <- import("LAPOP_merge_reduced.dta")
```

Esta base de datos solo contiene el grupo de variables necesarias para este análisis para todas las rondas y todos los países. El análisis para el gráfico 1 incluye a todos los países y el análisis del gráfico 2 incluye todas las rondas para Perú. Al momento de leer la base de datos en R, este programa importa las variables como numéricas. La variable “pais” se tiene que convertir en una variable de tipo “factor” y se tiene que etiquetar.

```
lapop$pais = as.factor(lapop$pais)
levels(lapop$pais) <- c("México", "Guatemala", "El Salvador", "Honduras",
  "Nicaragua", "Costa Rica", "Panamá", "Colombia",
  "Ecuador", "Bolivia", "Perú", "Paraguay", "Chile",
  "Uruguay", "Brasil", "Venezuela", "Argentina",
  "Rep. Dom.", "Haití", "Jamaica")
table(lapop$pais)
```

```
##
##      México  Guatemala El Salvador   Honduras   Nicaragua   Costa Rica
##      12476    12395    12488    12612    12607    12046
##      Panamá   Colombia   Ecuador    Bolivia    Perú      Paraguay
##      12455    12213    17991    21569    11668    9888
##      Chile    Uruguay    Brasil    Venezuela   Argentina   Rep. Dom.
##      11414    10319    11222    9068    8976    15047
##      Haití    Jamaica
##      10482    10629
```

## GRÁFICO 1

La autora indica que para este reporte se ha usado la pregunta: JC15A: ¿Cree usted que cuando el país enfrenta momentos muy difíciles, se justifica que el presidente del país cierre el Congreso y gobierne sin Congreso? + (1) Sí se justifica + (2) No se justifica Esta variable se analiza para la ronda 2018, por lo que se seleccionan estas observaciones en un nuevo dataframe “lapop18”.

```
lapop18 <- subset(lapop, wave==2018)
```

Esta base de datos ya cuenta con una variable recodificada llamada “jc15ar”, de tal manera que 1=100 y 2=0. Con el siguiente código se calcula la media de esta variable recodificada por cada país, tomando en cuenta la ponderación por país. Estos datos se guardan en un nuevo dataframe “tab.jc15ar”.

```
tab.jc15ar <- as.data.frame(compmeans(lapop18$jc15ar, lapop18$pais, lapop18$weight1500, plot=FALSE))
```

```
## Warning in compmeans(lapop18$jc15ar, lapop18$pais, lapop18$weight1500, plot =  
## FALSE): 11367 rows with missing values dropped
```

```
tab.jc15ar
```

```
##           Mean      N Std. Dev.  
## México      28.075052 1366  44.95304  
## Guatemala    22.797579 1398  41.96771  
## El Salvador  18.149965 1417  38.55677  
## Honduras     25.657427 1353  43.69039  
## Nicaragua    21.698113  719  41.24764  
## Costa Rica   17.261056  701  37.81803  
## Panamá       25.437416  715  43.58138  
## Colombia     16.546763  752  37.18493  
## Ecuador      23.956931  727  42.71145  
## Bolivia      27.272727  667  44.56960  
## Perú         58.923885  751  49.22996  
## Paraguay     27.167630  685  44.51489  
## Chile        21.279555  658  40.95954  
## Uruguay       9.221902  658  28.95548  
## Brasil       22.289784  716  41.64814  
## Venezuela    NA      NA    0.00000  
## Argentina    13.212796  706  33.88700  
## Rep. Dom.    24.305061 1388  42.90800  
## Haití        NA      NA    0.00000  
## Jamaica      23.959828  691  42.71475  
## Total        23.843688 16069 42.61409
```

En este dataframe, se cambian los nombres de las columnas, se agrega una nueva columna “pais”, se calcula el error estándar y el margen de error y se eliminan las filas sin datos y de total.

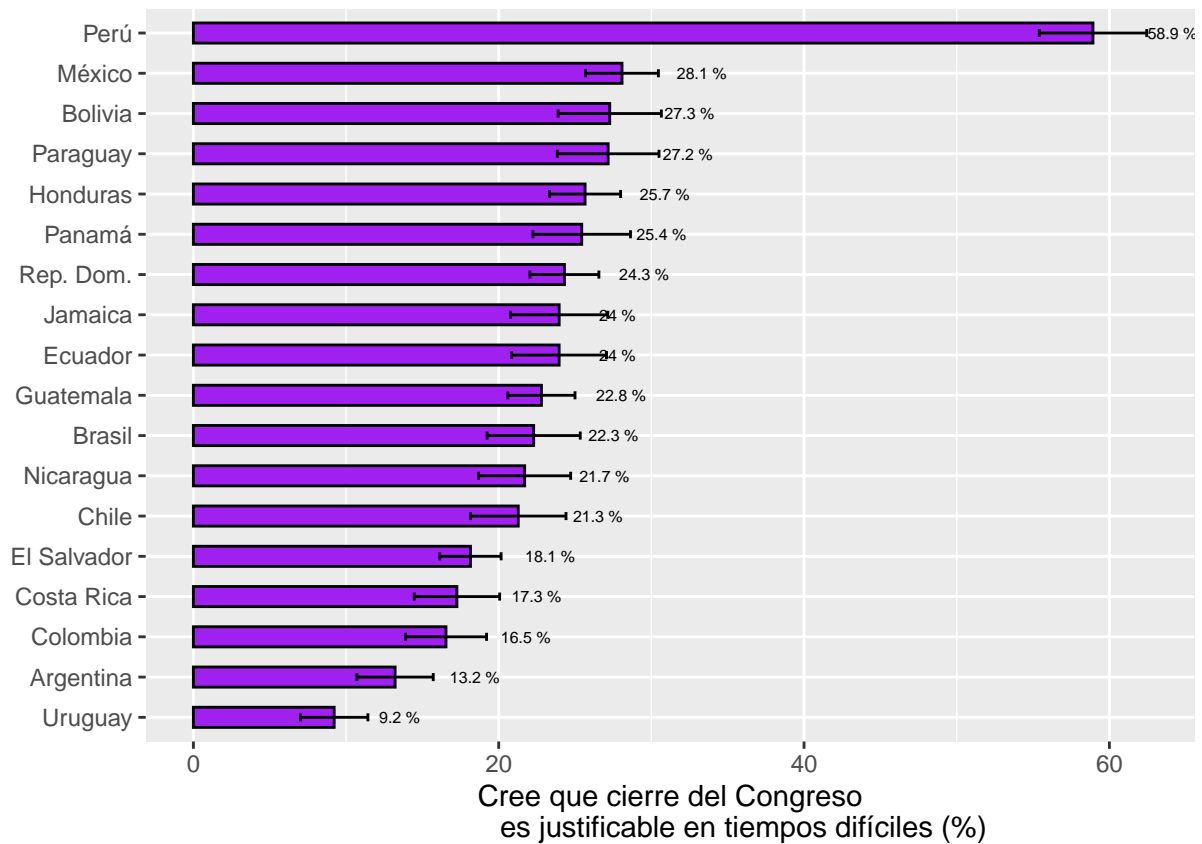
```
varnames <- c("media", "n", "sd")  
colnames(tab.jc15ar) <- varnames  
tab.jc15ar$pais <- row.names(tab.jc15ar)  
tab.jc15ar$err.st <- tab.jc15ar$sd/sqrt(tab.jc15ar$n)  
tab.jc15ar$ci <- tab.jc15ar$err.st*1.96  
tab.jc15ar <- tab.jc15ar[-21, ]  
tab.jc15ar <- tab.jc15ar[-16, ]  
tab.jc15ar <- tab.jc15ar[-18, ]  
tab.jc15ar
```

```
##           media      n      sd      pais  err.st      ci  
## México      28.075052 1366  44.95304 México 1.216280 2.383910  
## Guatemala    22.797579 1398  41.96771 Guatemala 1.122436 2.199975  
## El Salvador  18.149965 1417  38.55677 El Salvador 1.024273 2.007575  
## Honduras     25.657427 1353  43.69039 Honduras 1.187783 2.328054  
## Nicaragua    21.698113  719  41.24764 Nicaragua 1.538277 3.015024  
## Costa Rica   17.261056  701  37.81803 Costa Rica 1.428367 2.799600  
## Panamá       25.437416  715  43.58138 Panamá 1.629851 3.194508
```

## Colombia	16.546763	752	37.18493	Colombia	1.355995	2.657750
## Ecuador	23.956931	727	42.71145	Ecuador	1.584080	3.104797
## Bolivia	27.272727	667	44.56960	Bolivia	1.725742	3.382454
## Perú	58.923885	751	49.22996	Perú	1.796427	3.520997
## Paraguay	27.167630	685	44.51489	Paraguay	1.700826	3.333620
## Chile	21.279555	658	40.95954	Chile	1.596769	3.129668
## Uruguay	9.221902	658	28.95548	Uruguay	1.128802	2.212452
## Brasil	22.289784	716	41.64814	Brasil	1.556464	3.050669
## Argentina	13.212796	706	33.88700	Argentina	1.275354	2.499694
## Rep. Dom.	24.305061	1388	42.90800	Rep. Dom.	1.151711	2.257354
## Jamaica	23.959828	691	42.71475	Jamaica	1.624946	3.184893

Con estos datos en el dataframe se usa el siguiente código para producir el gráfico 1.

```
graf035_1 <- ggplot(tab.jc15ar, aes(x=reorder(pais, media), y=media)) +
  geom_bar(width=0.5, fill="purple", colour="black", stat="identity") +
  geom_errorbar(aes(ymin=media-ci, ymax=media+ci), width=0.2) +
  geom_text(aes(label=paste(round(media, 1), "%")), hjust=-1.1, size=2) +
  xlab("") + ylab("Cree que cierre del Congreso
                  es justificable en tiempos difíciles (%)") +
  coord_flip()
graf035_1
```



## GRÁFICO 2

Para producir el gráfico 2 se tiene que seleccionar las observaciones de Perú para los años 2010 en adelante.

```
peru <- subset(lapop, year>=2010 & pais=="Perú")
```

Con este dataframe se calcula el promedio de la variable “jc15ar” por año y se guarda en un nuevo dataframe “tab.peru”.

```
tab.peru <- as.data.frame(compmeans(peru$jc15ar, peru$year, peru$weight1500, plot=FALSE))
```

```
## Warning in compmeans(peru$jc15ar, peru$year, peru$weight1500, plot = FALSE):
```

```
## Warning: "peru$year" was converted into factor!
```

```
## Warning in compmeans(peru$jc15ar, peru$year, peru$weight1500, plot = FALSE):
```

```
## 1175 rows with missing values dropped
```

```
tab.peru
```

```
##           Mean      N Std. Dev.
## 2010  26.45390 1410  44.12439
## 2012  21.98839 1378  41.43182
## 2014  27.38351 1395  44.60853
## 2017  37.82106 1446  48.51083
## 2019  58.92388  751  49.22996
## Total 32.09307 6380  46.68708
```

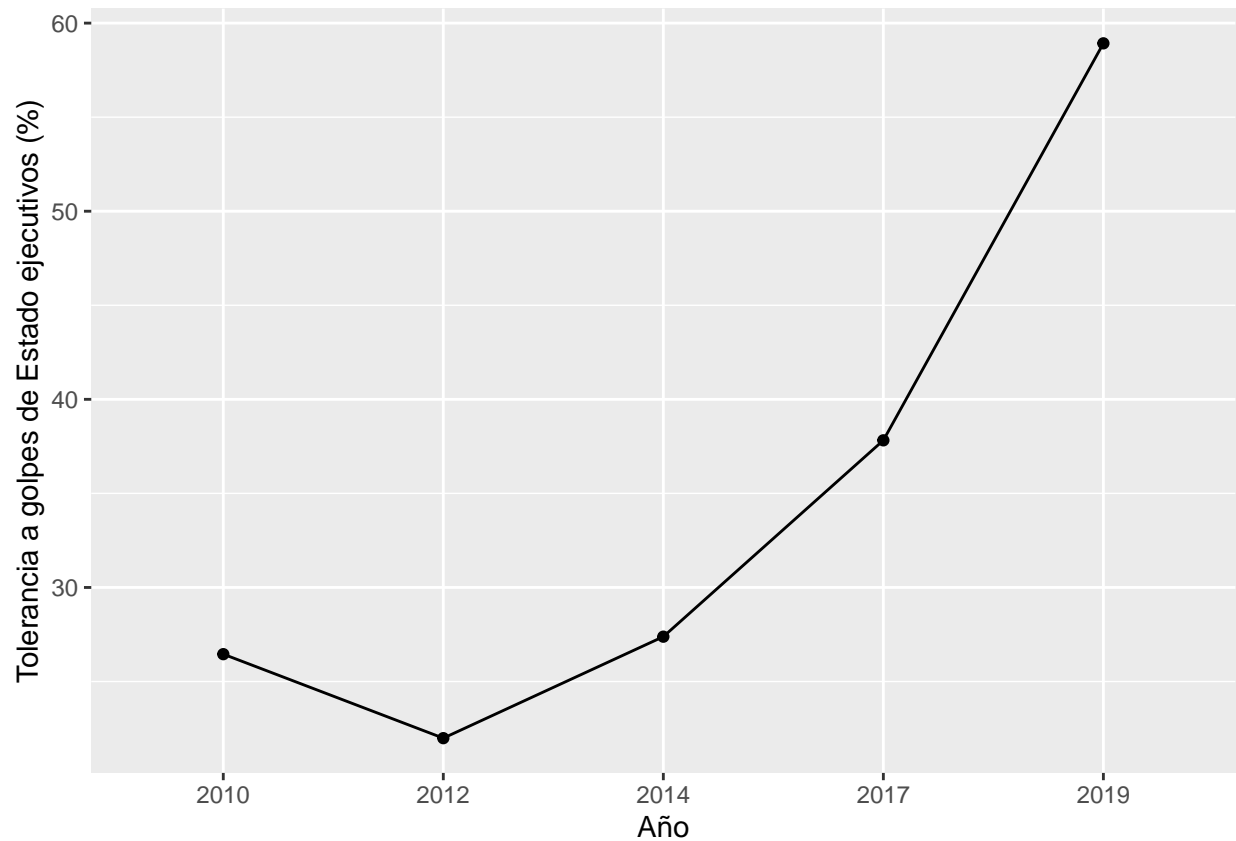
De la misma manera que en el gráfico 1, se tiene que adecuar este dataframe para tener los datos por año necesarios para producir el gráfico 2.

```
colnames(tab.peru) <- varnames
tab.peru$year <- row.names(tab.peru)
tab.peru$err.st <- tab.peru$sd/sqrt(tab.peru$n)
tab.peru$ci <- tab.peru$err.st*1.96
tab.peru <- tab.peru[-6, ]
tab.peru
```

```
##           media      n      sd year  err.st      ci
## 2010 26.45390 1410 44.12439 2010 1.175085 2.303166
## 2012 21.98839 1378 41.43182 2012 1.116116 2.187587
## 2014 27.38351 1395 44.60853 2014 1.194348 2.340922
## 2017 37.82106 1446 48.51083 2017 1.275718 2.500407
## 2019 58.92388  751 49.22996 2019 1.796427 3.520997
```

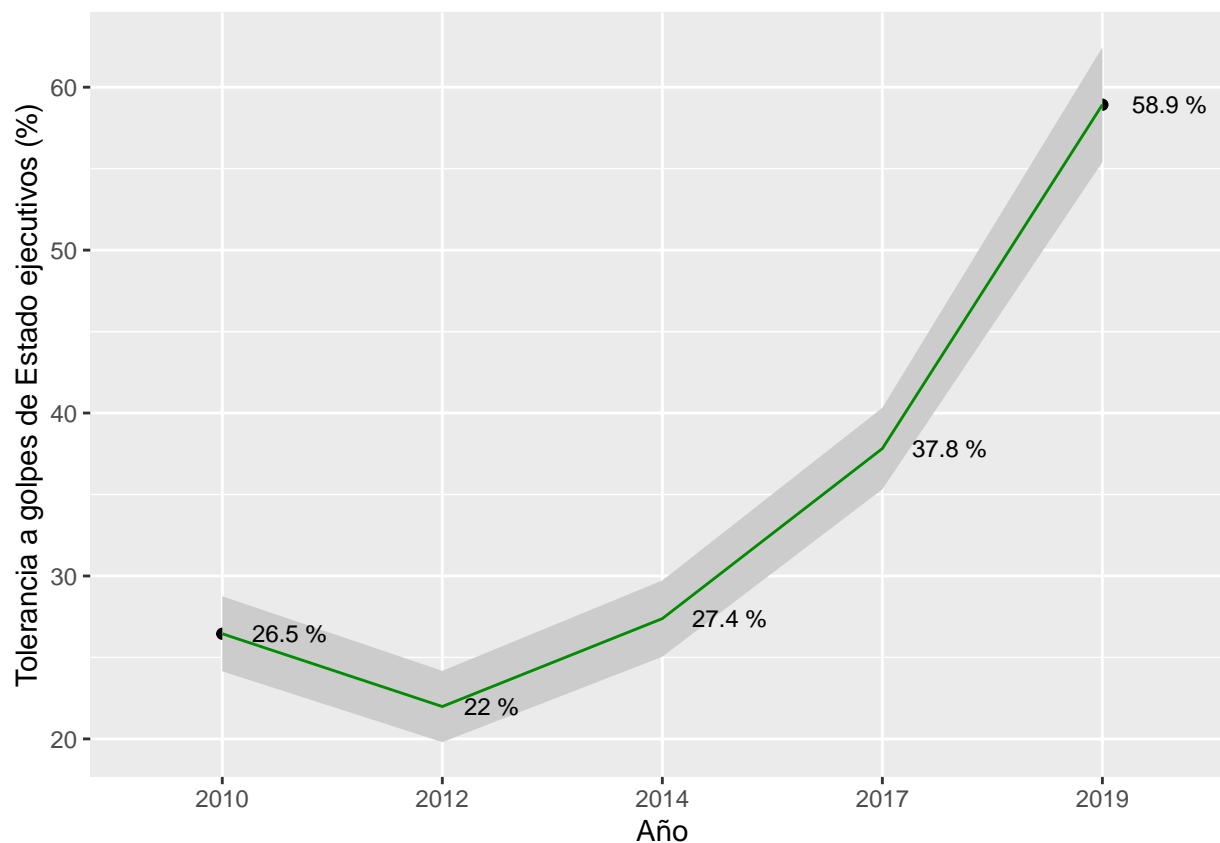
Primero se crea el gráfico de líneas que muestra el cambio en el promedio de esta variable por año para Perú.

```
graf2 <- ggplot(tab.peru, aes(x=year, y=media, group=1)) +
  geom_line() +
  geom_point() +
  ylab("Tolerancia a golpes de Estado ejecutivos (%)") +
  xlab("Año")
graf2
```



A este gráfico se le agrega la franja con el intervalo de confianza.

```
graf035_2 <- graf2 + geom_ribbon(aes(ymin=media-ci,
                                   ymax=media+ci),
                                linetype=1,
                                fill="grey80", outline.type="upper") +
  geom_line(aes(y=media), colour="green4") +
  geom_text(aes(label=paste(round(media, 1), "%")),
            hjust=-0.4, size=3)
graf035_2
```



De esta manera se han replicado los gráficos presentados en este Brief 035 de la Serie Perspectivas usando los datos del Barómetro de las Américas por LAPOP.

AmericasBarometer  
Barómetro de las Américas

**LAPOP**

ESCUELA DE  
GOBIERNO Y  
POLÍTICAS PÚBLICAS



**PUCP**