

# Análisis EPH con datos de panel

Economía Laboral

---

Miriam Malament

UCEMA

# Análisis EPH

Cargamos los paquetes y la panel de datos a utilizar

```
# PAQUETES
library(tidyverse)
library(data.table)
library(gt)
library(scales)
library(ggplot2)
library(eph)
```

```
#Importamos el panel de las EPH a utilizar
library(readxl)
panel ← read_excel("panel_eph.xlsx")
```

Vamos a estar trabajando con un panel de datos generados con los datos del primer trimestre de la Encuesta Permanente de Hogares para el 2004, 2007, 2011, 2015, 2017, 2019, 2022. Por cuestiones de comodidad, ya se encuentran seleccionadas las variables relevantes: 'CODUSU', 'NRO\_HOGAR', 'ANO4', 'TRIMESTRE', 'ESTADO', 'REGION', 'PONDERA', 'CH04', 'CH03', 'CH06', 'CH07', 'NIVEL\_ED'.

# Principales tasas del mercado laboral

Vamos a tomar el panel de datos con las EPH de 2004, 2007, 2011, 2015, 2017, 2019 y 2022 para analizar la trayectoria de las principales tasas del mercado laboral:

- Tasa de actividad como PEA/Población
- Tasa de empleo como Ocupados/Población
- Tasa de desocupación como Descuados/PEA

Asimismo, no solamente nos interesa la trayectoria de las tasas en sí, sino también cómo fue cambiando a lo largo de las últimas dos décadas según el rango etario, género, región, estado civil, nivel educativo y relación de parentesco.

Algunas de las preguntas que vamos a querer contestar:

- ¿Cómo cambió la trayectoria de las tasas en el tiempo?
- ¿Qué tanto varían las tasas entre géneros?
- ¿Cómo varían las tasas entre regiones?
- ¿A mayor nivel educativo mayor tasa de empleo?

# ¿Cómo realizar los filtrados?

- La función `case_when` permite definir una variable, la cual toma un valor particular para cada condición establecida.
- Se usa así: `case_when(condicion1 ~ "Valor1",condicion2 ~ "Valor2",condicion3 ~ "Valor3")`
- Lo que queremos hacer es agregar una columna a la panel de datos en la que se le asigne a cada observación su rango etario. Aquellos individuos entre 25 y 34 años entrarán dentro del rango etario "25-34", a los que tengan más de 35 pero menos de 44 se les asignará "35-44" y así sucesivamente.
- Notar que aquellos individuos con menos de 25 o más de 65 se les asignará NA.

```
panel <- panel %>%  
  mutate(rango_etario = case_when(CH06 ≥ 25 & CH06 ≤ 34 ~ "25-34",  
                                   CH06 ≥ 35 & CH06 ≤ 44 ~ "35-44",  
                                   CH06 ≥ 45 & CH06 ≤ 54 ~ "45-54",  
                                   CH06 ≥ 55 & CH06 ≤ 64 ~ "55-64"))
```

Estos filtros ya fueron realizados como he mencionado anteriormente.

# Tasa de Actividad

---

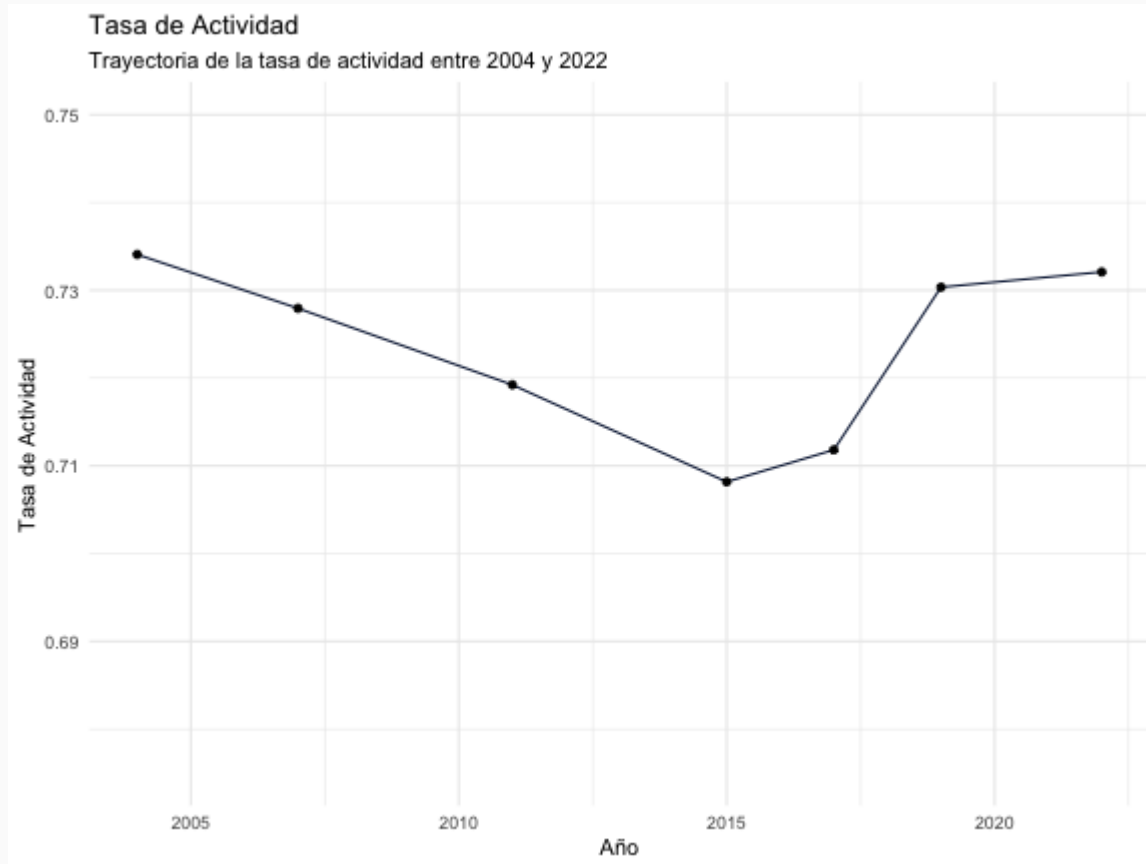
# Tasa de Actividad

- Primero, vamos a calcular la tasa de actividad por año.
- Para eso, agrupamos por año y calculamos los ocupados, desocupados, la población económicamente activa (PEA) y la tasa de actividad como la fracción entre PEA y población.
- Recordemos que la tasa de actividad se calcula como el porcentaje entre la población económicamente activa y la población total de referencia..
- Importante destacar que el ponderador a utilizar es PONDERA.

```
actividad_año ← panel %>%  
  group_by(ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población)
```

# Tasa de Actividad

Graficamos la trayectoria de la tasa de actividad:



# Tasa de Actividad

Año 2022

Vayamos paso a paso y arranquemos analizando un año en particular: 2022.

Vamos a filtrar por rango etario y género

```
actividad_edad <- panel %>% filter(ANO4== 2022)%>%  
  group_by(rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            "Tasa Actividad" = PEA/Población) %>% na.omit()
```

```
actividad_edad_genero <- panel %>% filter(ANO4==2022)%>%  
  group_by(rango_etario, genero) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit()
```



# Tasa de Actividad

Año 2022

Ahora filtramos por rango etario, región y educación

```
actividad_region_edad <- panel %>% filter(ANO4=2022)%>%  
  group_by(region, rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO = 1]),  
            Desocupados = sum(PONDERA[ESTADO = 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población)  
  
actividad_educacion_edad <- panel %>% filter(ANO4=2022)%>%  
  group_by(educacion, rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO = 1]),  
            Desocupados = sum(PONDERA[ESTADO = 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit()
```

# Tasa de Actividad

Año 2022

Tasa de Actividad				
Por rango etario, género, región y nivel educativo				
	25-34	35-44	45-54	55-64
Nivel Educativo				
Primario completo o menos	66.17%	76.21%	72.53%	54.46%
Secundario completo o menos	79.14%	83.70%	80.47%	62.45%
Terciario incompleto o más	83.18%	91.44%	90.66%	71.31%
Región				
Cuyo	83.79%	86.46%	83.97%	63.68%
GBA	52.55%	80.23%	85.83%	82.93%
Nordeste	63.05%	52.42%	74.55%	82.72%
Noroeste	76.49%	52.48%	43.99%	76.52%
Pampeana	84.77%	80.35%	62.02%	44.54%
Datos correspondientes al primer trimestre del 2022				

# Tasa de Actividad

## ANÁLISIS POR RANGO ETARIO Y GÉNERO

```
actividad_edad_genero_panel <- panel %>%  
  group_by(rango_etario, genero, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población)
```

## `summarise()` has grouped output by 'rango\_etario', 'genero'. You can override  
## using the `.groups` argument.

- Como no estamos interesados en aquellos individuos menores a 25 ni mayores a 65, aplicamos `na.omit()`

```
actividad_edad_genero_panel <- actividad_edad_genero_panel %>% na.omit()
```

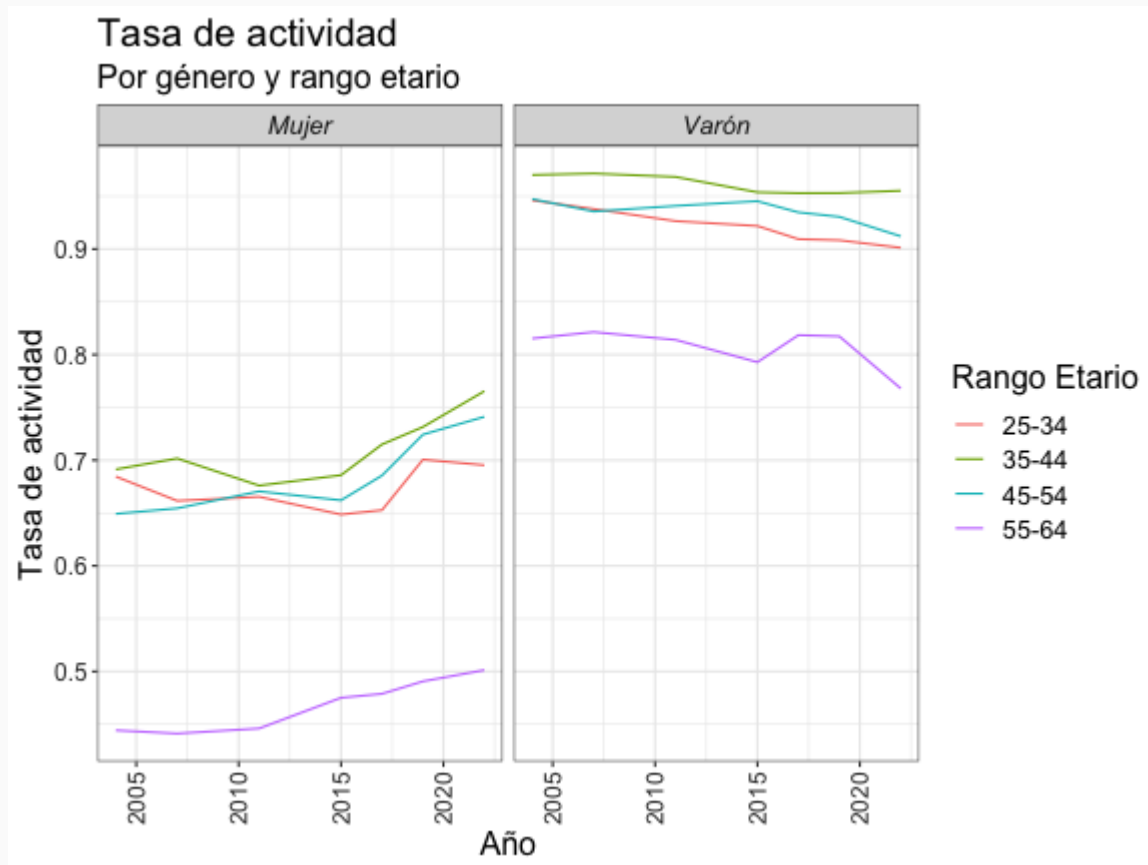
# Tasa de Actividad

Tabla de la tasa de actividad:

<b>Tasa de Actividad</b>							
Por rango etario y año							
	2004	2007	2011	2015	2017	2019	2022
25-34	68.46%	66.2%	66.5%	64.9%	65.3%	70.1%	69.5%
35-44	94.61%	93.8%	92.6%	92.2%	90.9%	90.8%	90.1%
45-54	69.14%	70.2%	67.6%	68.6%	71.5%	73.2%	76.5%
55-64	97.01%	97.2%	96.8%	95.4%	95.3%	95.3%	95.5%
Datos correspondientes al 1T del 2004, 2007, 2011, 2015, 2017, 2019, 2022							

# Tasa de Actividad

## ANÁLISIS POR RANGO ETARIO Y GÉNERO



# Tasa de Actividad

## ANÁLISIS POR REGIÓN

```
actividad_region_panel <- panel %>%  
  group_by(region, AN04) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población)
```

# Tasa de Actividad

## ANÁLISIS POR REGIÓN

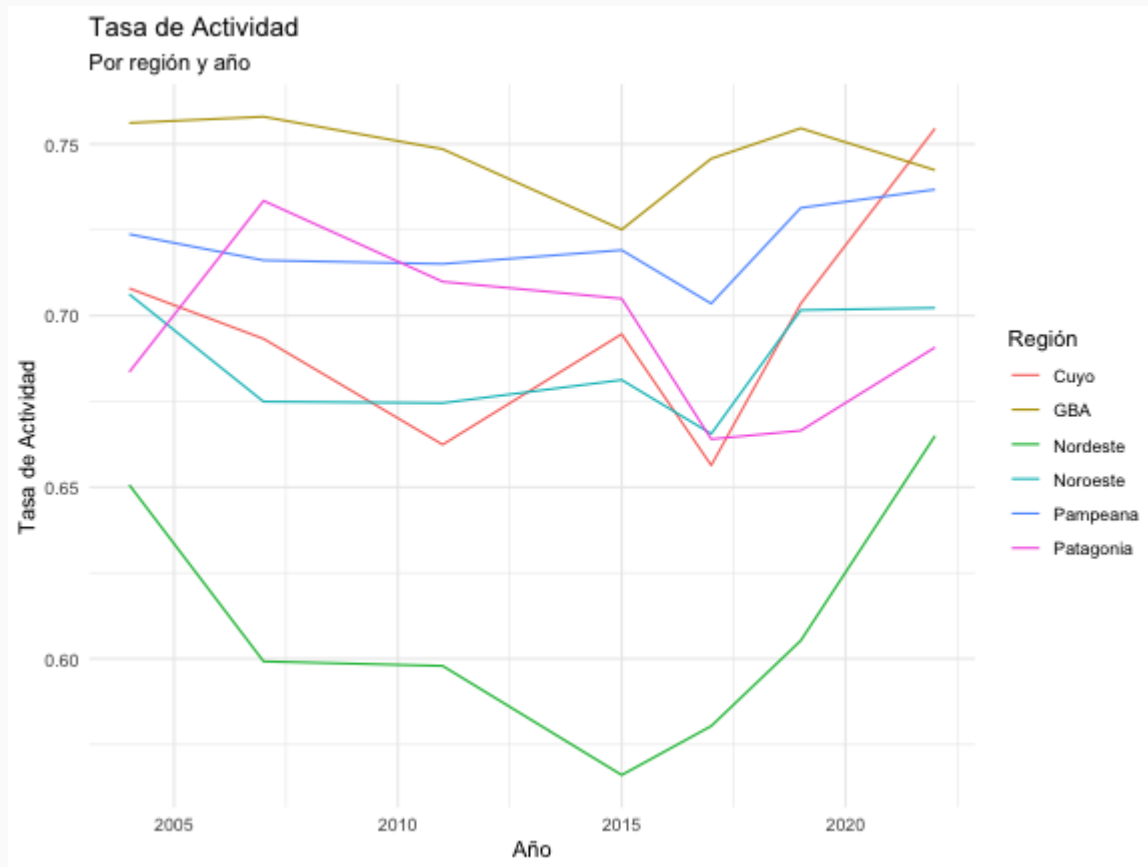
Podemos hacer una tabla:

Tasa de Actividad						
Por región						
	Cuyo	GBA	Nordeste	Noroeste	Pampeana	Patagonia
2004	70.80%	75.62%	65.07%	70.623%	72.37%	68.36%
2007	69.32%	75.80%	59.92%	67.493%	71.61%	73.35%
2011	66.24%	74.85%	59.79%	67.452%	71.51%	70.99%
2015	69.46%	72.51%	56.61%	68.126%	71.91%	70.49%
2017	65.64%	74.58%	58.04%	66.555%	70.35%	66.40%
2019	70.35%	75.46%	60.53%	70.160%	73.14%	66.65%
2022	75.45%	74.24%	66.50%	70.224%	73.68%	69.07%
Datos correspondientes al 1T del 2004, 2007, 2011, 2015, 2017, 2019, 2022						

# Tasa de Actividad

## ANÁLISIS POR REGIÓN

Graficamos





# Tasa de Actividad

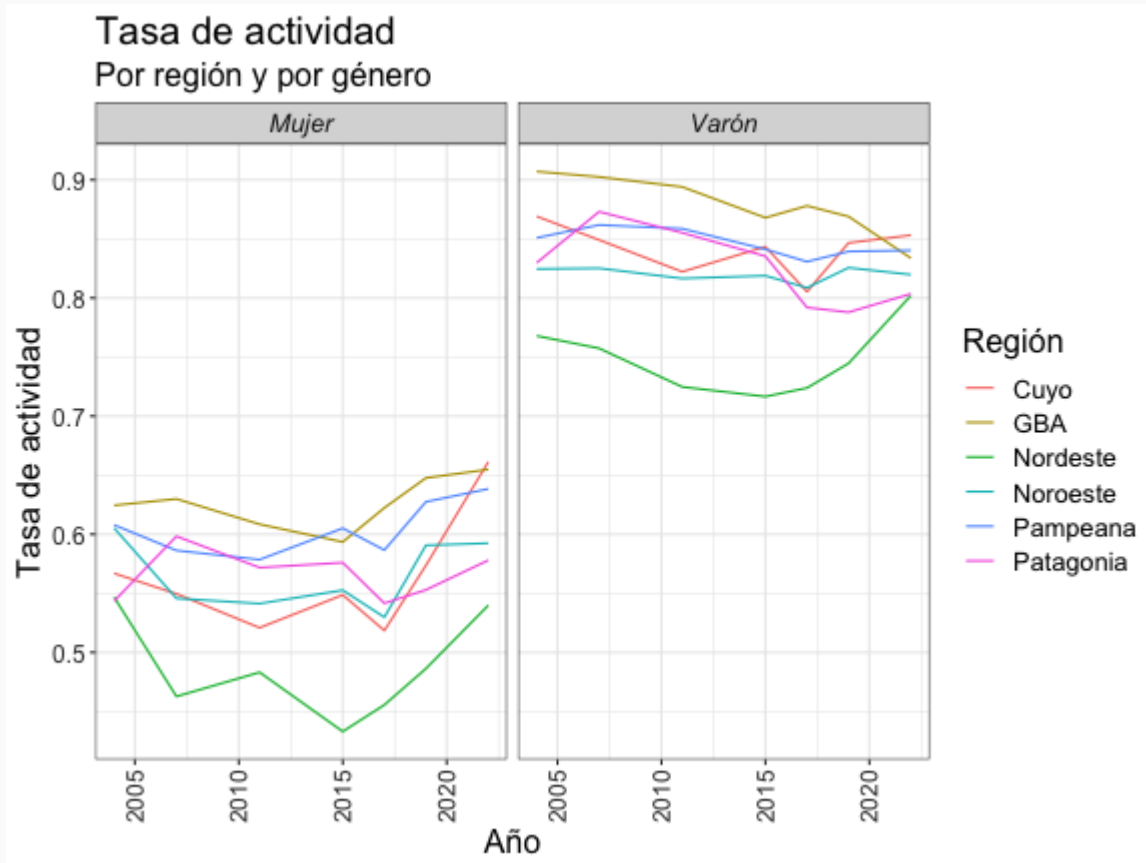
## ANÁLISIS POR REGIÓN DIFERENCIANDO POR GÉNERO

```
actividad_region_genero_panel <- panel %>%  
  group_by(region, genero, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit()
```

## `summarise()` has grouped output by 'region', 'genero'. You can override using  
## the `.groups` argument.

# Tasa de Actividad

## ANÁLISIS POR REGIÓN DIFERENCIANDO POR GÉNERO



# Tasa de Actividad

## ANÁLISIS POR ESTADO CIVIL (mujeres)

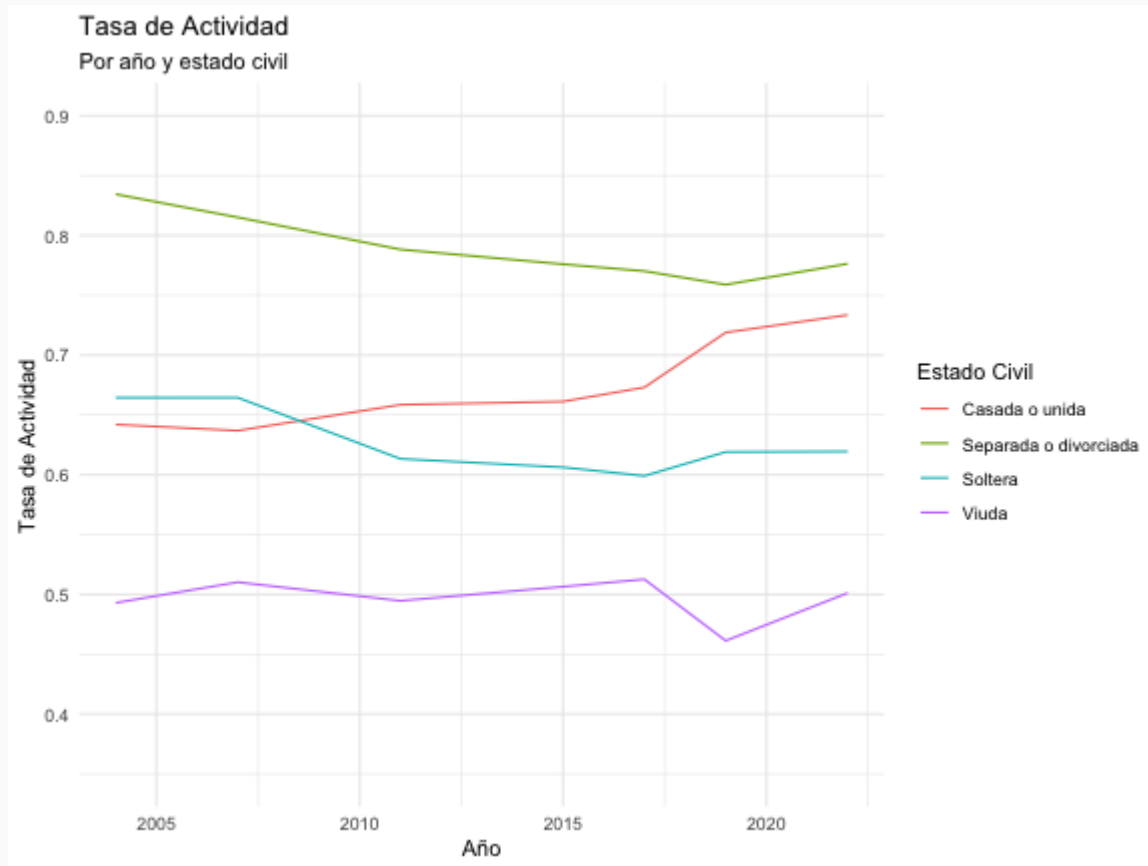
```
actividad_estado_civil_panel <- panel %>%  
  group_by(estado_civil, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit()
```

```
## `summarise()` has grouped output by 'estado_civil'. You can override using the  
## `.groups` argument.
```

# Tasa de Actividad

## ANÁLISIS POR ESTADO CIVIL (mujeres)

Graficamos:



# Tasa de Actividad

## ANÁLISIS POR NIVEL EDUCATIVO PARA VARONES

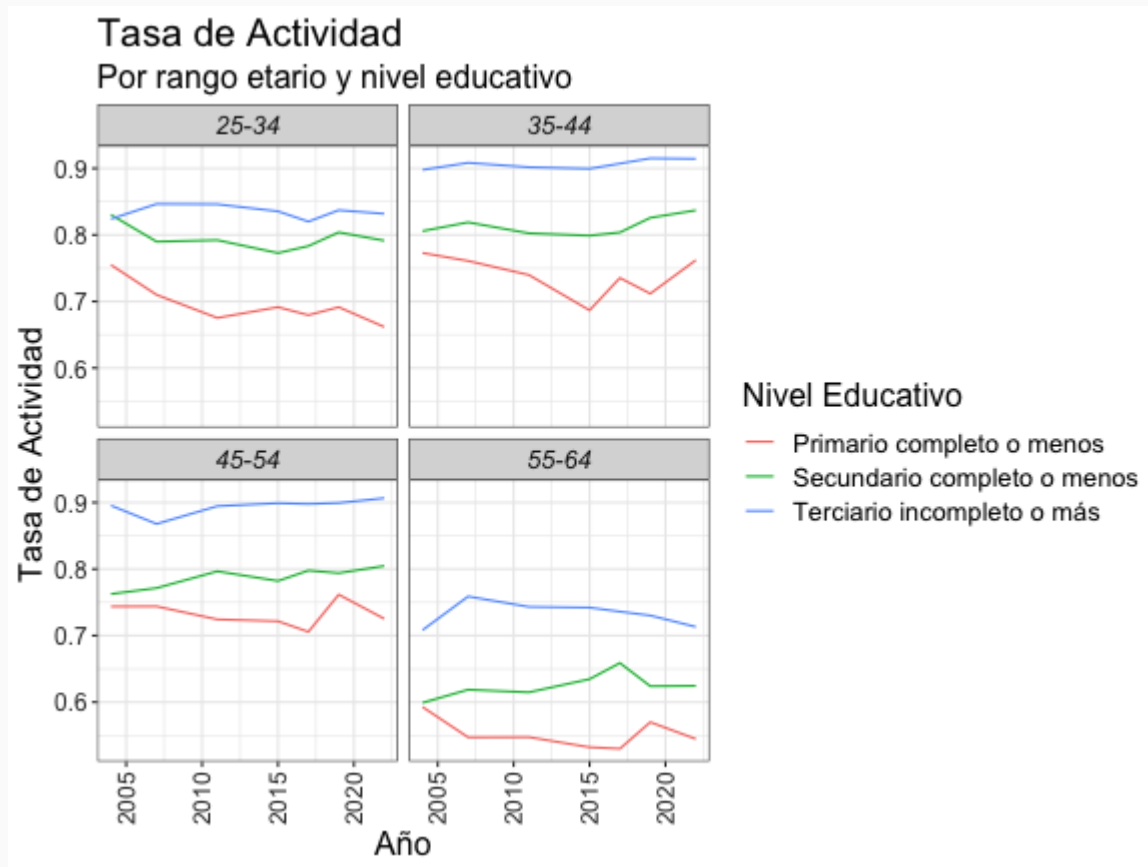
```
actividad_educacion_edad_hombres_panel <- panel %>%  
  group_by(educacion, rango_etario, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit() #No quiero considerar a las
```

```
## `summarise()` has grouped output by 'educacion', 'rango_etario'. You can  
## override using the `.groups` argument.
```

# Tasa de Actividad

## ANÁLISIS POR NIVEL EDUCATIVO PARA VARONES

Graficamos:



# Tasa de Actividad

## ANÁLISIS POR NIVEL EDUCATIVO PARA MUJERES

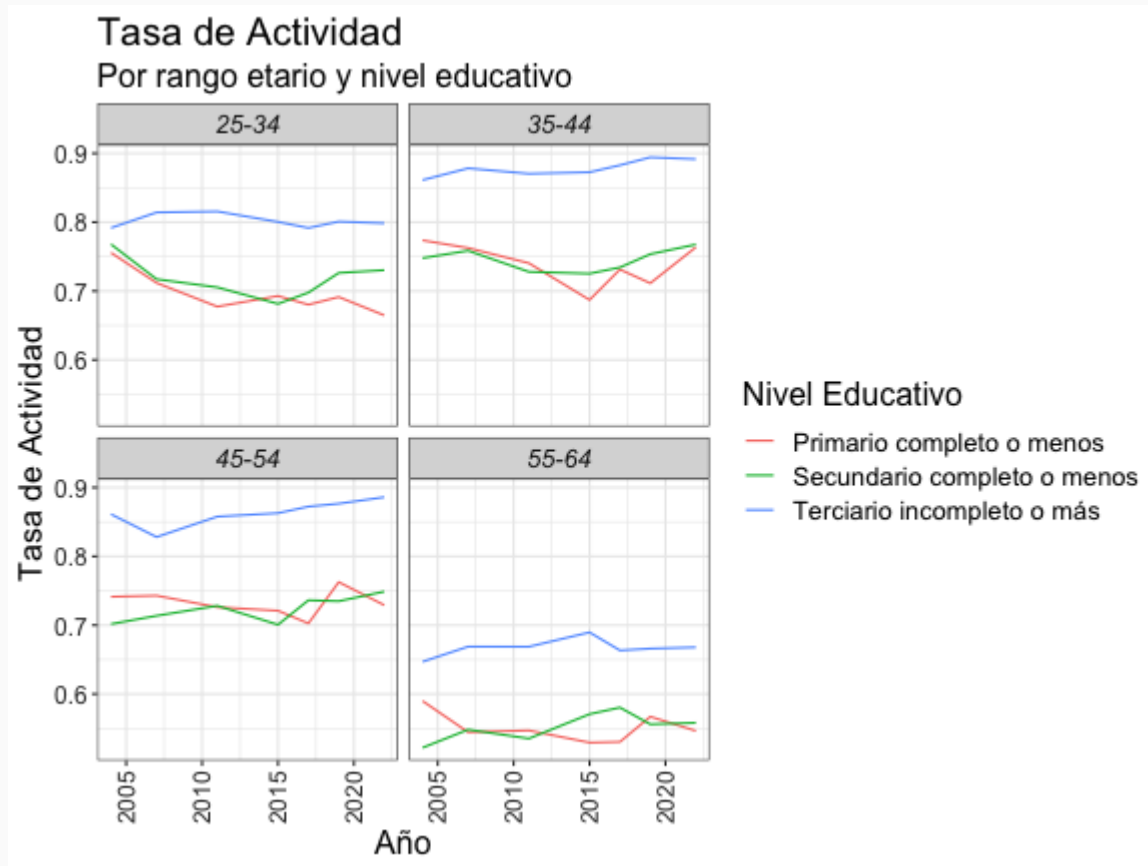
```
actividad_educacion_edad_mujeres_panel <- panel %>%  
  group_by(educacion_m, rango_etario, AN04) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Actividad = PEA/Población) %>% na.omit
```

```
## `summarise()` has grouped output by 'educacion_m', 'rango_etario'. You can  
## override using the `.groups` argument.
```

# Tasa de Actividad

## ANÁLISIS POR NIVEL EDUCATIVO PARA MUJERES

Graficamos:





# Tasa de Empleo

---

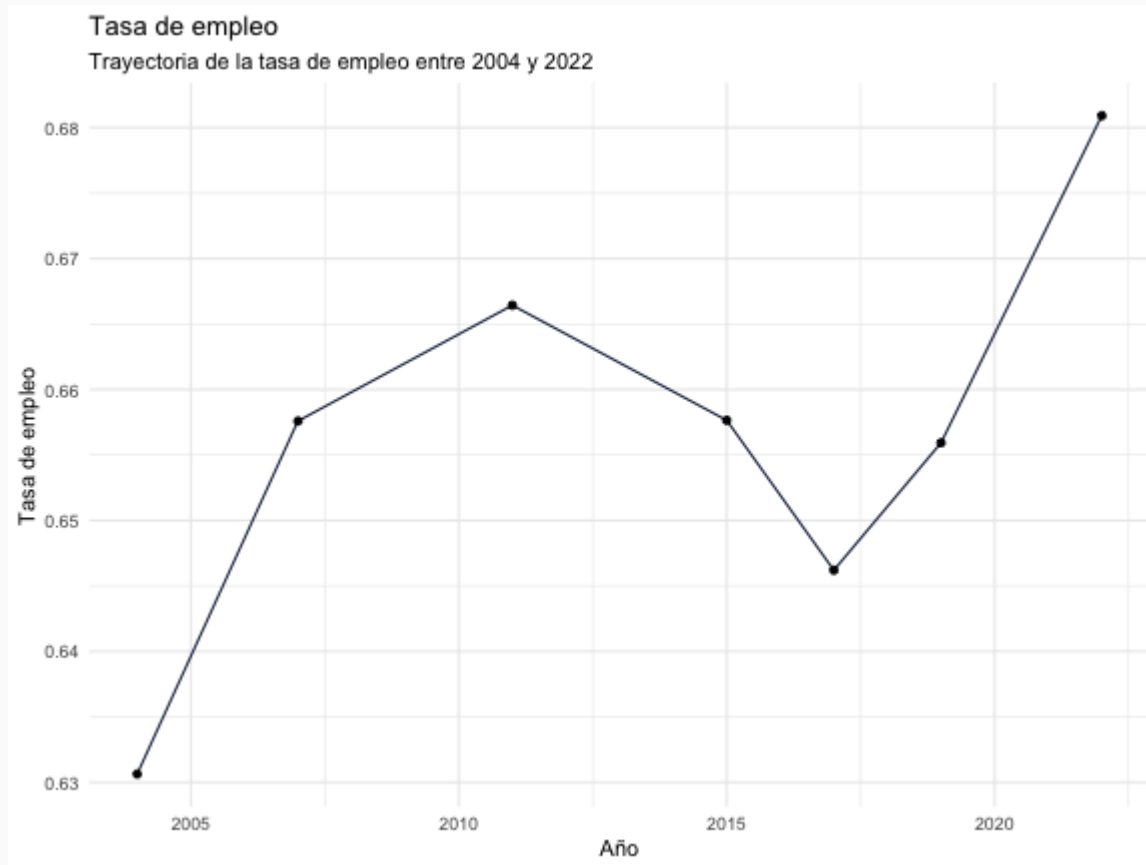
# Tasa de Empleo

- La tasa de actividad está compuesta por la proporción de la población que forma parte de la fuerza laboral. Eso quiere decir que toma en cuenta a las personas ocupadas y desocupadas (las que no tienen trabajo, pero se encuentran buscándolo activamente). Ahora, nos vamos a concentrar en la tasa de empleo (Ocupados/Población).
- Para eso, agrupamos por año y calculamos los ocupados, desocupados, la población económicamente activa (PEA) y la tasa de empleo como la fracción entre Ocupados y Población.
- Importante destacar que el ponderador a utilizar es PONDERA.

```
empleo_año ← panel %>%  
  group_by(ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población)
```

# Tasa de Empleo

Graficamos la trayectoria de la tasa de empleo:



# Tasa de Empleo

Año 2022

Vayamos paso a paso y arranquemos analizando un año en particular: 2022.

Vamos a filtrar por rango etario y género

```
empleo_edad <- panel %>% filter(AN04== 2022)%>%  
  group_by(rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            "Tasa Actividad" = PEA/Población) %>% na.omit()
```

```
empleo_edad_genero <- panel %>% filter(AN04==2022)%>%  
  group_by(rango_etario, genero) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit()
```

# Tasa de Empleo

Año 2022

Ahora filtramos por rango etario, región y educación

```
empleo_region_edad <- panel %>% filter(ANO4=2022)%>%  
  group_by(region, rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO = 1]),  
            Desocupados = sum(PONDERA[ESTADO = 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población)  
  
empleo_educacion_edad <- panel %>% filter(ANO4=2022)%>%  
  group_by(educacion, rango_etario) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO = 1]),  
            Desocupados = sum(PONDERA[ESTADO = 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit()
```

# Tasa de Empleo

Año 2022

<b>Tasa de Actividad</b>				
<b>Por rango etario, género, región y nivel educativo</b>				
	25-34	35-44	45-54	55-64
<b>Nivel Educativo</b>				
Primario completo o menos	62.49%	72.00%	70.00%	50.54%
Secundario completo o menos	73.00%	77.47%	76.80%	59.99%
Terciario incompleto o más	78.62%	88.41%	86.45%	70.23%
<b>Región</b>				
Cuyo	78.42%	82.90%	82.40%	61.25%
GBA	45.50%	75.46%	80.83%	78.53%
Nordeste	60.55%	42.22%	70.30%	80.17%
Noroeste	74.76%	51.62%	37.63%	69.90%
Pampeana	80.07%	78.29%	60.11%	38.11%
Datos correspondientes al primer trimestre del 2022				

# Tasa de Empleo

## ANÁLISIS POR RANGO ETARIO Y GÉNERO

```
empleo_edad_genero_panel <- panel %>%  
  group_by(rango_etario, genero, AN04) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población)
```

## `summarise()` has grouped output by 'rango\_etario', 'genero'. You can override  
## using the `.groups` argument.

- Como no estamos interesados en aquellos individuos menores a 25 ni mayores a 65, aplicamos `na.omit()`

```
empleo_edad_genero_panel <- empleo_edad_genero_panel %>% na.omit()
```

# Tasa de Empleo

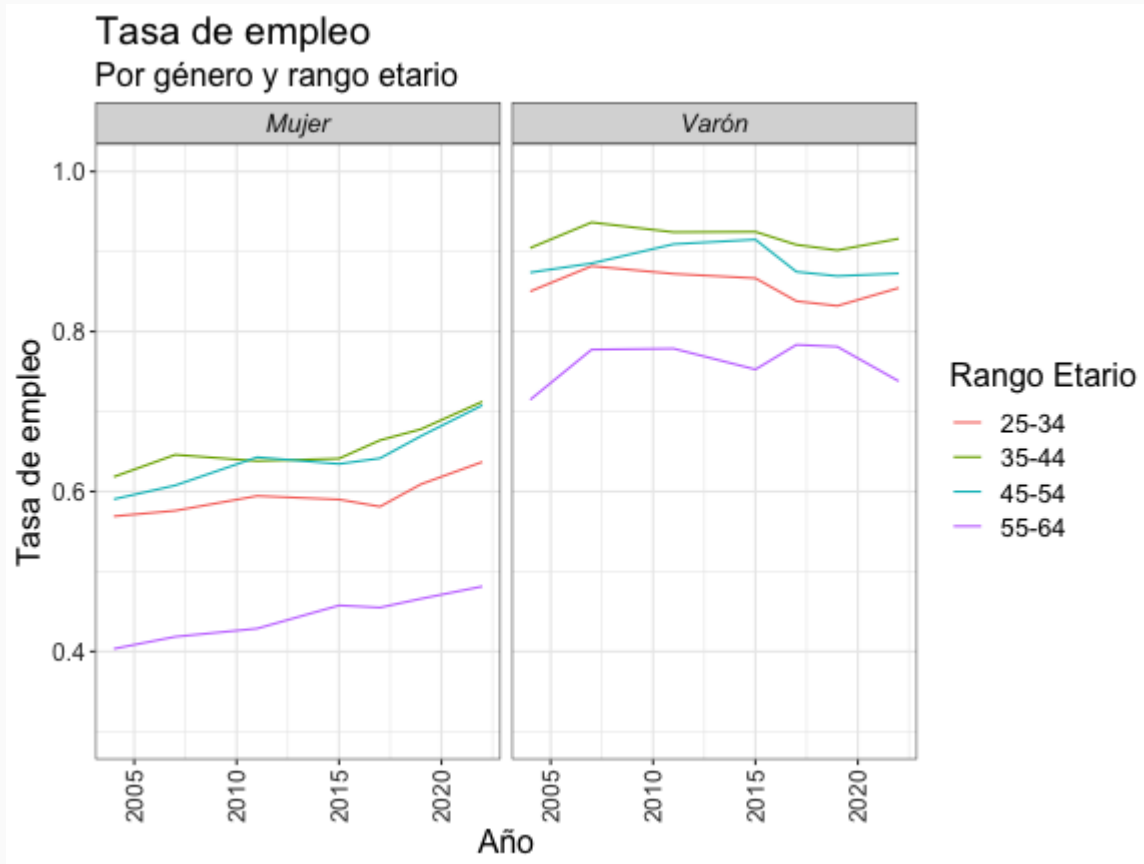
Tabla de la tasa de empleo:

<b>Tasa de Empleo</b>							
Por rango etario y año							
	2004	2007	2011	2015	2017	2019	2022
25-34	56.9%	57.6%	59.4%	59.0%	58.1%	60.9%	63.7%
35-44	85.0%	88.2%	87.2%	86.7%	83.8%	83.2%	85.4%
45-54	61.8%	64.6%	63.8%	64.1%	66.4%	67.8%	71.2%
55-64	90.4%	93.6%	92.4%	92.5%	90.8%	90.2%	91.6%
Datos correspondientes al 1T del 2004, 2007, 2011, 2015, 2017, 2019, 2022							



# Tasa de Empleo

## ANÁLISIS POR RANGO ETARIO Y GÉNERO



# Tasa de Empleo

## ANÁLISIS POR REGIÓN

```
empleo_region_panel <- panel %>%  
  group_by(region, AN04) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población)
```

# Tasa de Empleo

## ANÁLISIS POR REGIÓN

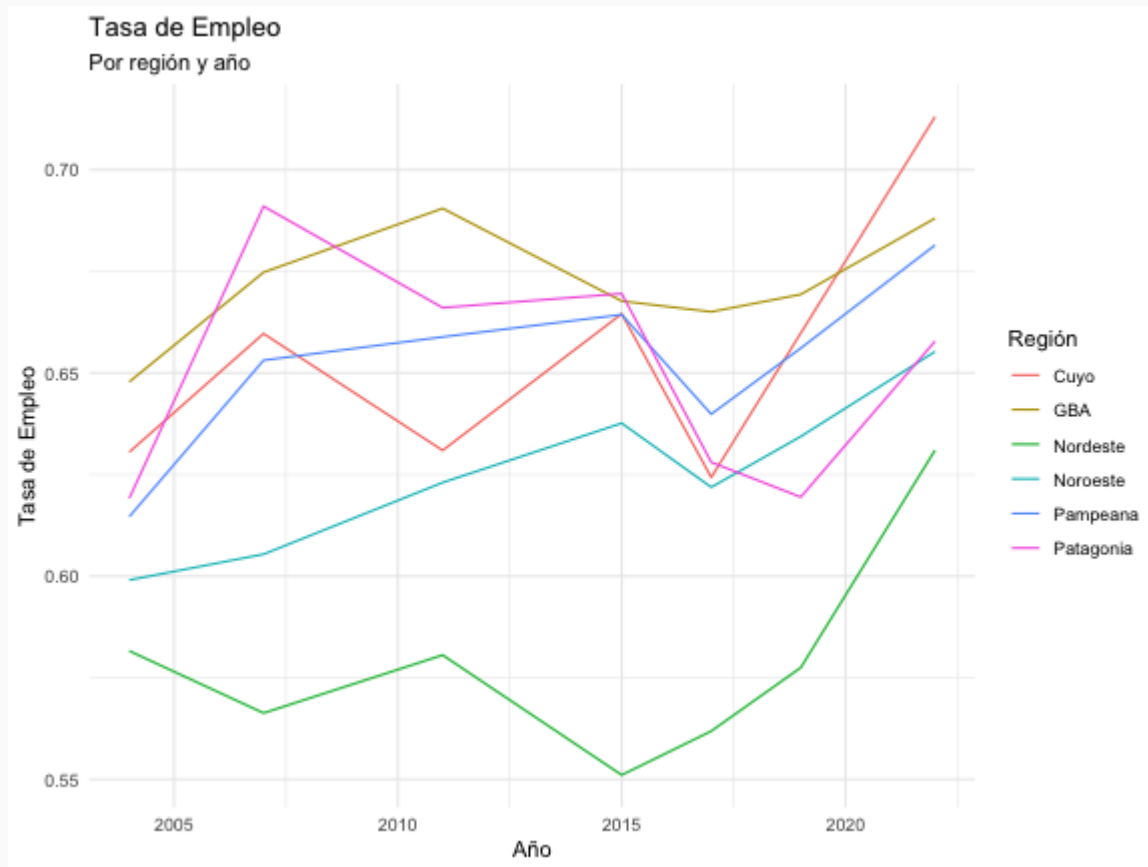
Podemos hacer una tabla:

Tasa de Empleo						
Por región						
	Cuyo	GBA	Nordeste	Noroeste	Pampeana	Patagonia
2004	63.051%	64.78%	58.16%	59.90%	61.46%	61.914%
2007	65.967%	67.48%	56.63%	60.54%	65.31%	69.099%
2011	63.093%	69.05%	58.06%	62.30%	65.89%	66.605%
2015	66.453%	66.77%	55.11%	63.76%	66.44%	66.952%
2017	62.430%	66.50%	56.19%	62.19%	63.98%	62.802%
2019	65.979%	66.93%	57.75%	63.44%	65.61%	61.945%
2022	71.304%	68.81%	63.10%	65.53%	68.14%	65.781%
Datos correspondientes al 1T del 2004, 2007, 2011, 2015, 2017, 2019, 2022						

# Tasa de Empleo

## ANÁLISIS POR REGIÓN

Graficamos



# Tasa de Empleo

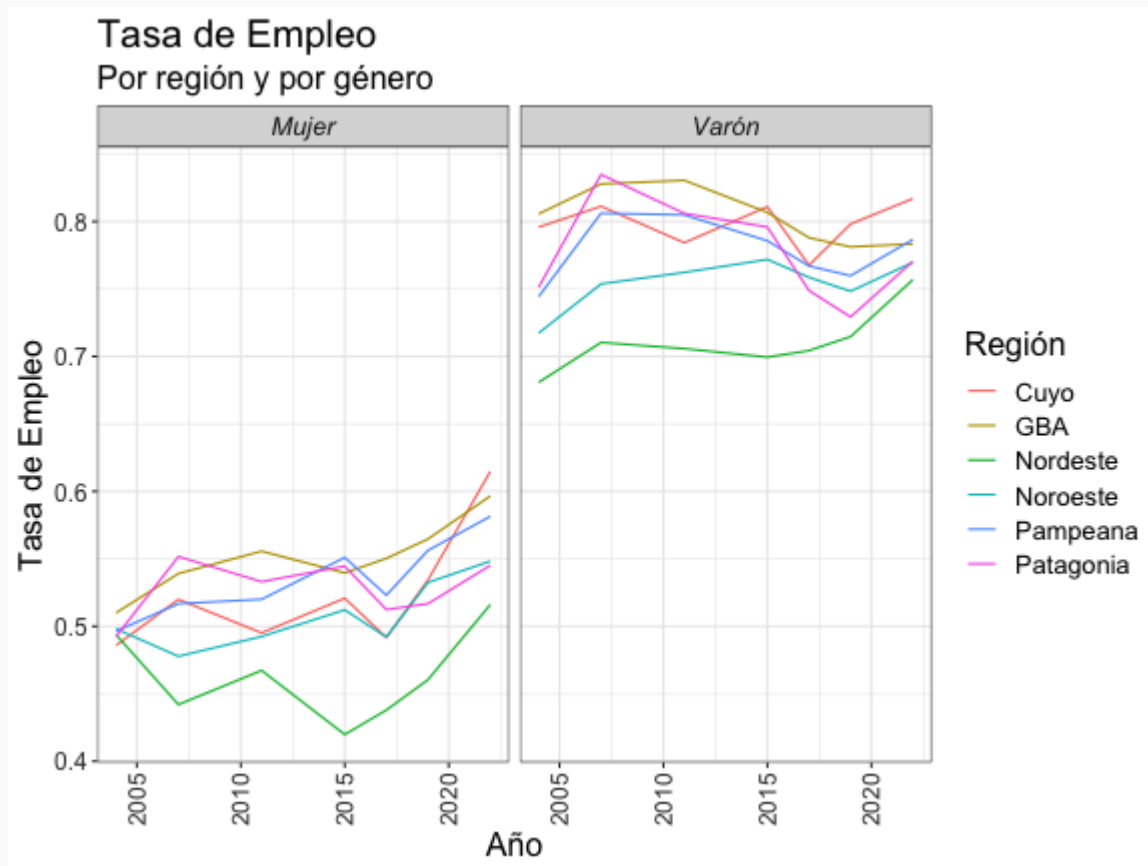
## ANÁLISIS POR REGIÓN DIFERENCIANDO POR GÉNERO

```
empleo_region_genero_panel <- panel %>%  
  group_by(region, genero, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit()
```

## `summarise()` has grouped output by 'region', 'genero'. You can override using  
## the `.groups` argument.

# Tasa de Empleo

## ANÁLISIS POR REGIÓN DIFERENCIANDO POR GÉNERO



# Tasa de Empleo

## ANÁLISIS POR ESTADO CIVIL (mujeres)

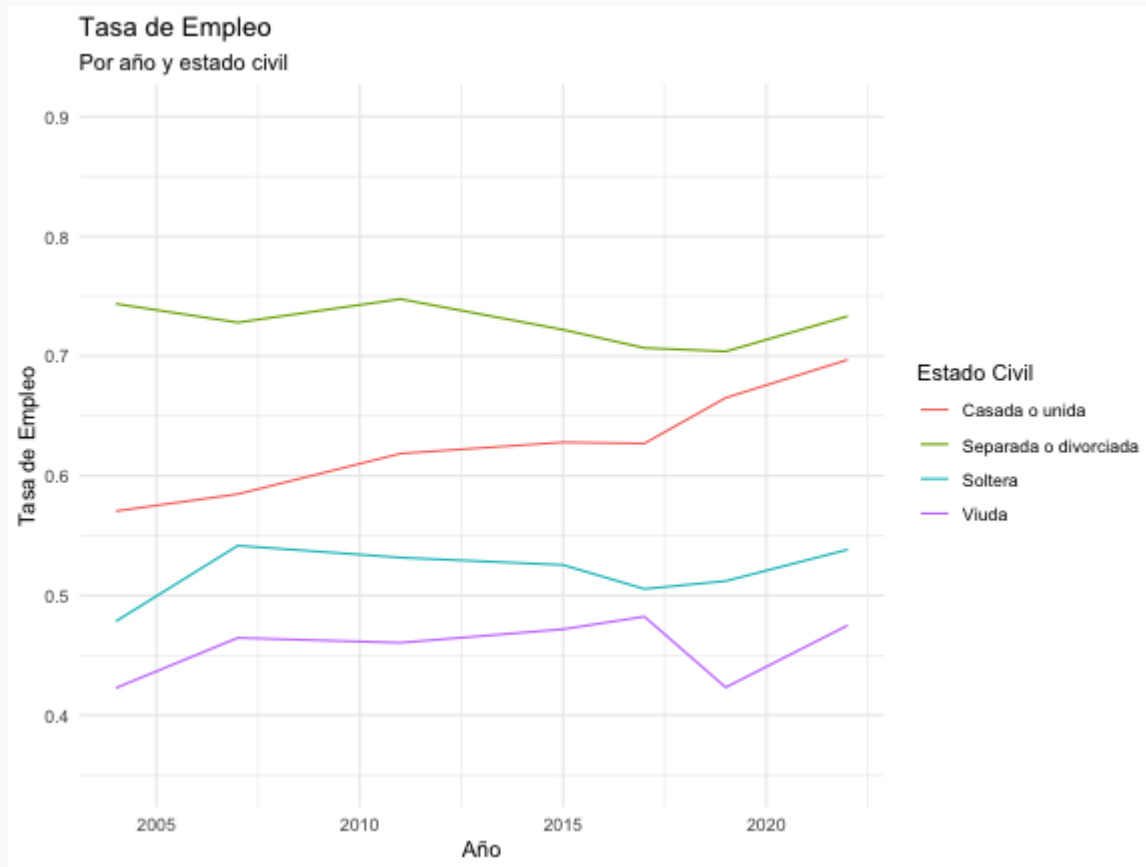
```
empleo_estado_civil_panel <- panel %>%  
  group_by(estado_civil, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit()
```

```
## `summarise()` has grouped output by 'estado_civil'. You can override using the  
## `.groups` argument.
```

# Tasa de Empleo

## ANÁLISIS POR ESTADO CIVIL (mujeres)

Graficamos:





# Tasa de Empleo

## ANÁLISIS POR NIVEL EDUCATIVO PARA VARONES

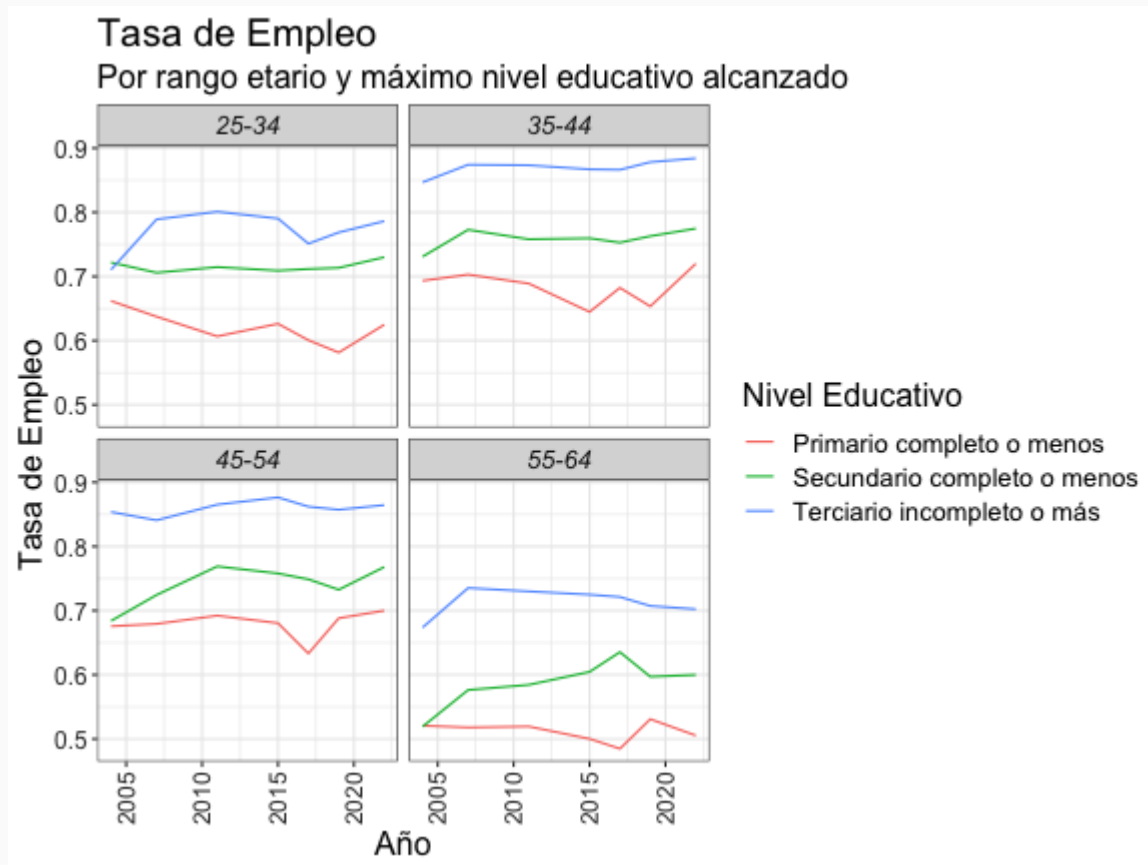
```
empleo_educacion_edad_hombres_panel <- panel %>%  
  group_by(educacion, rango_etario, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit() #No quiero considerar a la
```

```
## `summarise()` has grouped output by 'educacion', 'rango_etario'. You can  
## override using the `.groups` argument.
```

# Tasa de Empleo

## ANÁLISIS POR NIVEL EDUCATIVO PARA VARONES

Graficamos:



# Tasa de Empleo

## ANÁLISIS POR NIVEL EDUCATIVO PARA MUJERES

```
empleo_educacion_edad_mujeres_panel <- panel %>%  
  group_by(educacion_m, rango_etario, ANO4) %>%  
  summarise(Población = sum(PONDERA),  
            Ocupados = sum(PONDERA[ESTADO == 1]),  
            Desocupados = sum(PONDERA[ESTADO == 2]),  
            PEA = Ocupados + Desocupados,  
            Tasa_Empleo = Ocupados/Población) %>% na.omit
```

```
## `summarise()` has grouped output by 'educacion_m', 'rango_etario'. You can  
## override using the `.groups` argument.
```

# Tasa de Empleo

## ANÁLISIS POR NIVEL EDUCATIVO PARA MUJERES

Graficamos:

