

# Parte I: Familiarizandonos con la base EPH y limpieza

1-

El INDEC identifica la desocupación mediante la Encuesta Permanente de Hogares (EPH). Esta encuesta, recaba información sobre la población adulta, incluyendo datos sobre su situación laboral, características demográficas y socioeconómicas. Dentro de la EPH, se considera desocupado a quien no ha trabajado en el período de referencia, está disponible para trabajar y ha buscado activamente empleo.

2-

b) Para el año 2024:

Las variables con más valores faltantes son:

PP3E\_TOT (Ingreso de la ocupación principal): 665 valores faltantes (46.77%)

PP04D\_COD (Código de ocupación): 665 valores faltantes (46.77%)

PP04B\_COD (Código de actividad): 665 valores faltantes (46.77%)

P47T (Ingreso total): 6 valores faltantes (0.42%)

Para el año 2004:

No se encontraron valores faltantes en ninguna de las variables seleccionadas.

Las variables relacionadas con ingresos y ocupación en 2024 (PP3E\_TOT, PP04D\_COD, PP04B\_COD) tienen un alto porcentaje de valores faltantes (46.77%), lo que podría afectar significativamente los análisis relacionados con ingresos y ocupación.

El ingreso total (P47T) en 2024 tiene muy pocos valores faltantes (solo 6 casos, 0.42%).

Es notable que 2004 no presenta valores faltantes, lo que sugiere que la base de ese año podría haber tenido un tratamiento previo de los datos faltantes o un mejor proceso de recolección de datos.

Las variables demográficas básicas (como CH04, CH03, etc.) están completas en ambos años.

c)

Valores sin sentido encontrados:

+ Variable P21 (Horas trabajadas):

En 2004: valores entre 0 y 10000 horas

En 2024:

- Se encuentran valores negativos (mínimo -9)
- Las horas trabajadas no pueden ser negativas ni exceder valores razonables (una semana tiene 168 horas)

+ Variables de Ingreso (ITF, IPCF, P47T, PP3E\_TOT):

Para P47T (Ingreso total):

2024: presenta valores negativos (mínimo -9)

Los valores negativos: se eliminaron porque:

Las horas trabajadas no pueden ser negativas

Los ingresos no pueden ser negativos

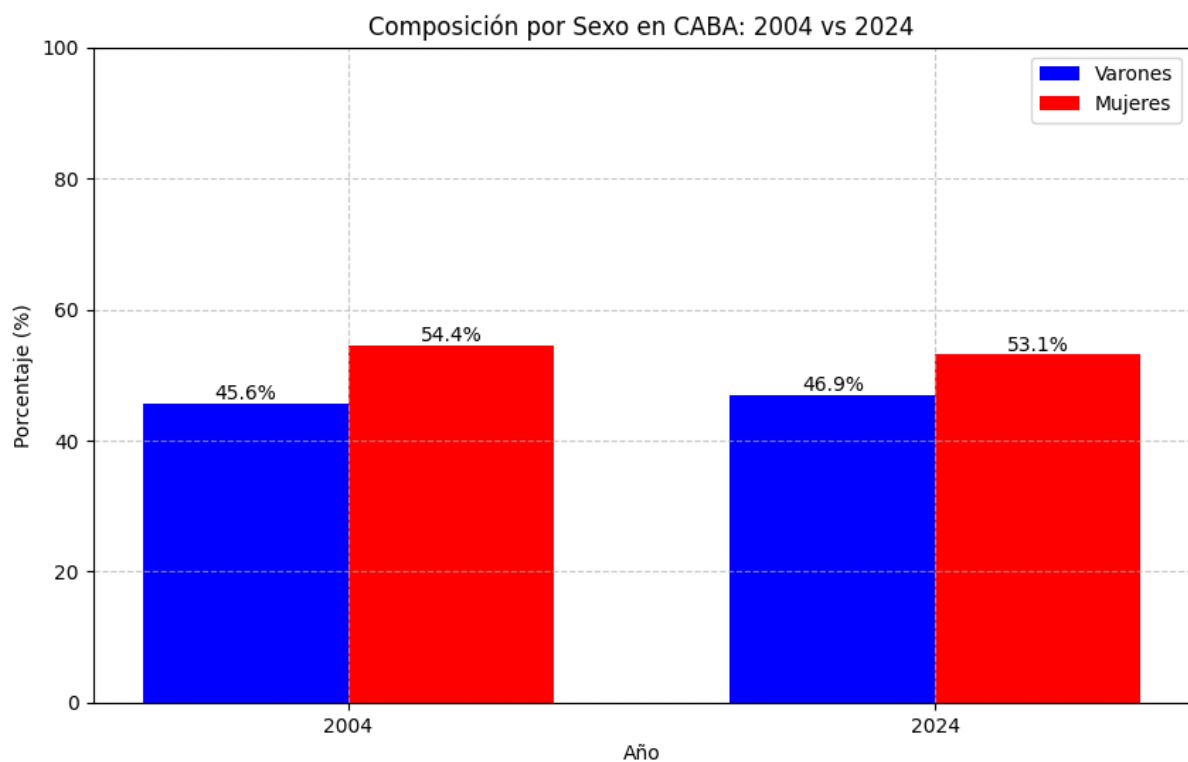
Horas trabajadas excesivas: Se limitaron a 168 horas semanales porque:

Es el máximo físicamente posible (7 días x 24 horas)

Valores superiores son claramente errores de registro

## Parte II: Primer Análisis Exploratorio

a)



Año 2004: Diferencia de 8.8 puntos porcentuales a favor de las mujeres

Año 2024: Diferencia de 6.2 puntos porcentuales a favor de las mujeres

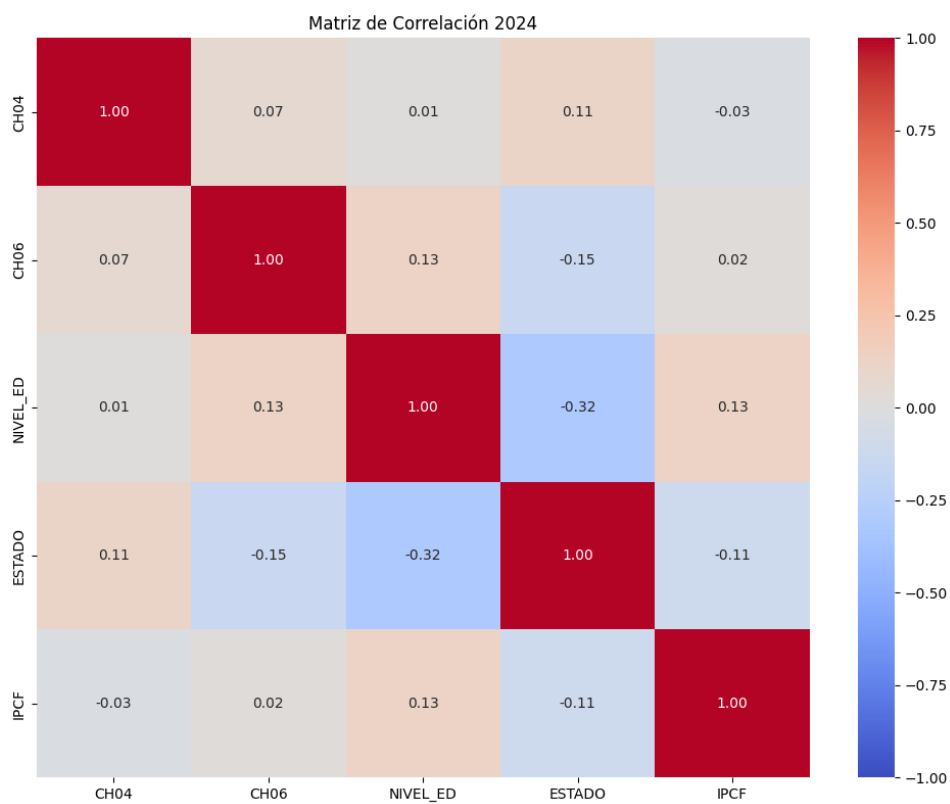
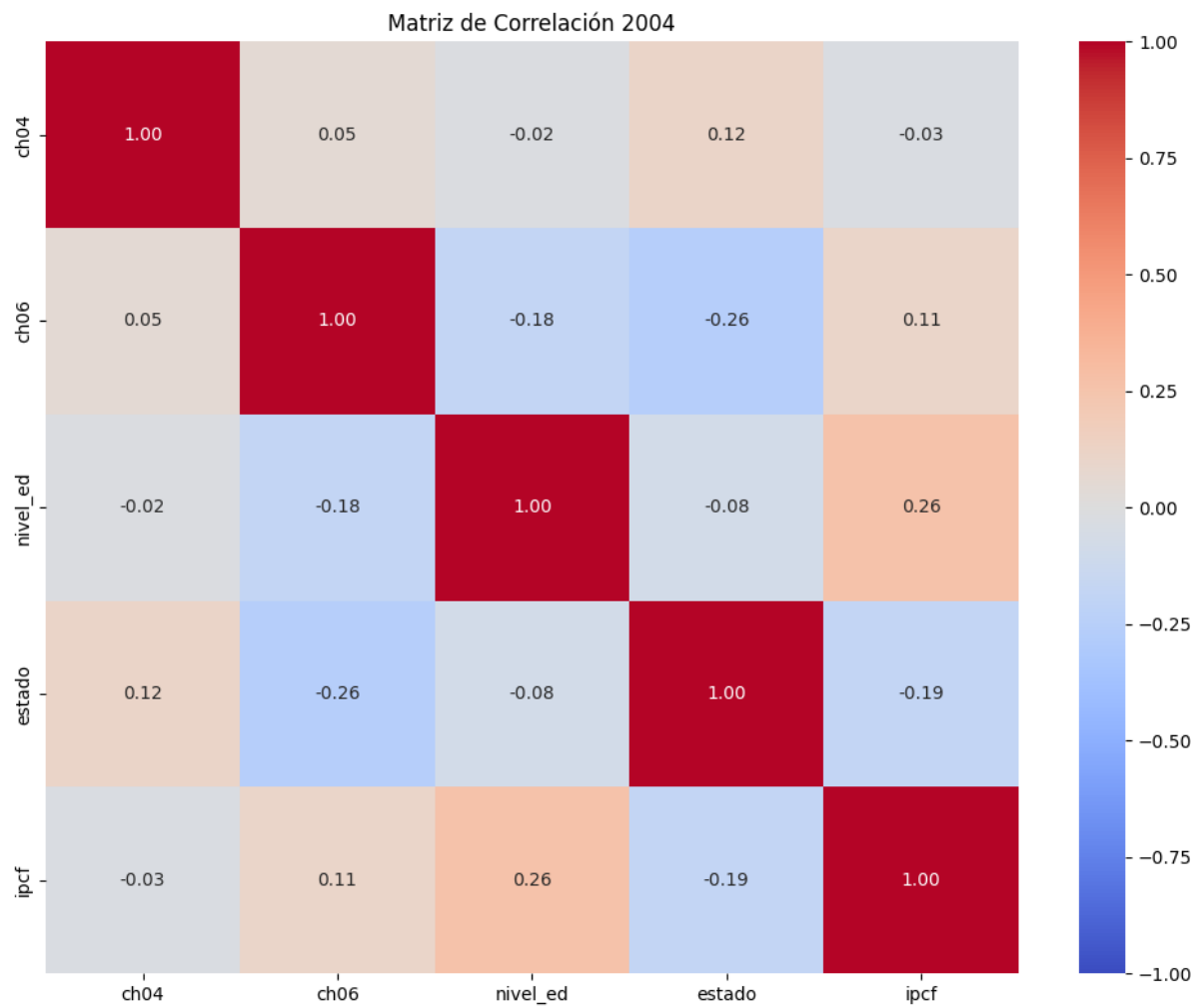
Cambios entre 2004 y 2024:

La proporción de mujeres disminuyó levemente (1.3 puntos porcentuales)

Se observa una tendencia hacia un mayor equilibrio entre géneros

En conclusión, en ambos años se mantiene una mayoría femenina y la brecha de género se redujo ligeramente en el período de 20 años

b)



## Parte III: Conociendo a los ocupados y desocupados