

Trabajo Práctico 1 – Taller de programación

Maestría: Economía aplicada

Profesor: Noelia Romero

Alumnos: Shanthal Chavarria, Edgar Julián Pineda Toro, Gustavo Horacio Romero

Enlace del repositorio: <https://github.com/shanthalchs/BigDataUBA-Grupo6>

1. Introducción

En este trabajo analizamos la información proveniente de la *Encuesta Permanente de Hogares (EPH)* correspondiente al primer trimestre de 2005 y 2025, para la región del **Gran Buenos Aires (GBA)**. El objetivo principal fue conocer la evolución de las características sociodemográficas, educativas y laborales de la población, así como las relaciones entre estas dimensiones. Adicionalmente, buscamos identificar patrones de cambio estructural entre ambos años mediante herramientas de programación en **Python**, empleando las bibliotecas *pandas*, *numpy*, *seaborn* y *matplotlib*.

2. Metodología

Cargamos las bases de microdatos de individuos de los años 2005 (formato .dta) y 2025 (formato .xlsx). Luego, filtramos las observaciones correspondientes al **Gran Buenos Aires** y unificamos ambas bases en un solo DataFrame. Para garantizar la consistencia de la información:

- Convertimos los nombres de las variables del año 2005 a mayúsculas.
- Estandarizamos los tipos de datos (numéricos y categóricos).
- Reemplazamos los valores **0** en variables categóricas que no admiten dicha categoría (por ejemplo, *CAT_INAC*, *CAT_OCUP*, *PP07H*, *PP04G*) por valores faltantes (NaN).
- Eliminamos observaciones con valores imposibles o faltantes en las 15 variables seleccionadas.

Con la base limpia, realizamos análisis descriptivos, exploratorios y una evaluación de correlaciones.

Adicionalmente, realizamos los diferentes tipos de unión con la base de hogares respectiva de cada año, obteniendo los siguientes resultados:

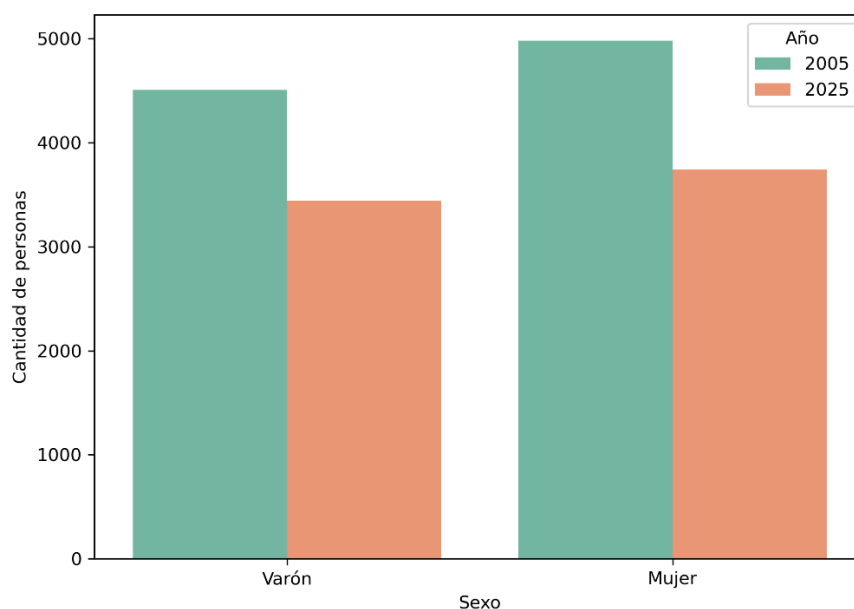
Tabla 1. Unión de las bases individuales con hogares – Gran Buenos Aires (2005 y 2025)

Tipo de unión	Nº filas	Nº Columnas	Total, NAs
Unión por intersección	16,665	21	4,879
Unión por izquierda	16,665	21	4,879
Unión por derecha	16,665	21	4,879
Unión conjunta	16,665	21	4,879

3. Resultados

En el Gráfico 1 se muestra la composición por sexo para los años 2005 y 2025 en el Gran Buenos Aires. En 2005, el **52.5 %** de la población era mujer (4.980 personas) y el **47.5 %** varón (4.504 personas). En 2025, los porcentajes se mantuvieron estables, con **52.1 %** mujeres (3.742) y **47.9 %** varones (3.439)

Gráfico 1. Composición por sexo en Gran Buenos Aires – EPH 2005 y 2025



Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

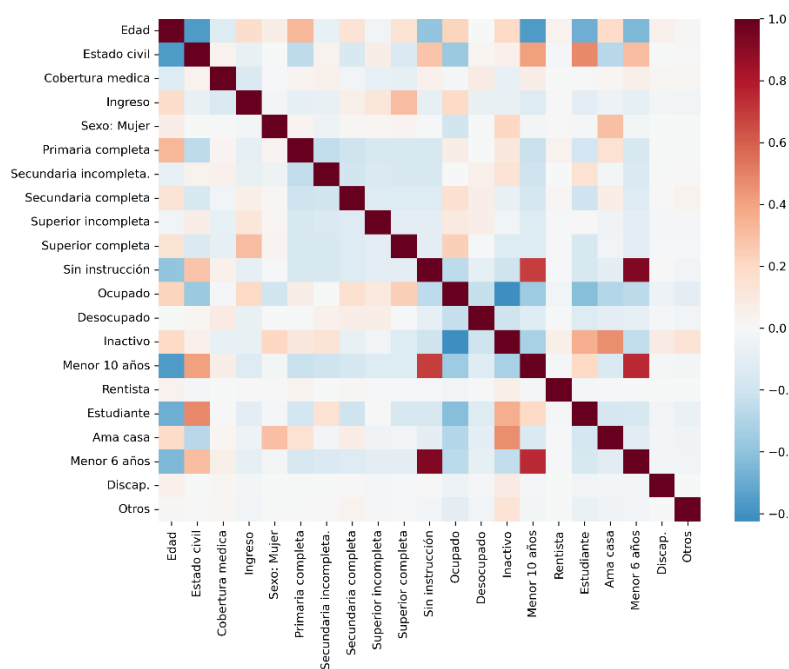
Para examinar las relaciones entre variables, construimos matrices de correlación para ambos años considerando: **CH04 (sexo)**, **CH06 (edad)**, **CH07**, **CH08**, **NIVEL_ED**, **ESTADO**, **CAT_INAC** e **IPCF**. Las variables categóricas fueron transformadas en variables dicotómicas (dummies) para el cálculo de correlaciones de Pearson.

En el *Gráfico 2* y *Gráfico 3* se presentan las matrices de correlación de 2005 y 2025 respectivamente.

Entre los principales hallazgos:

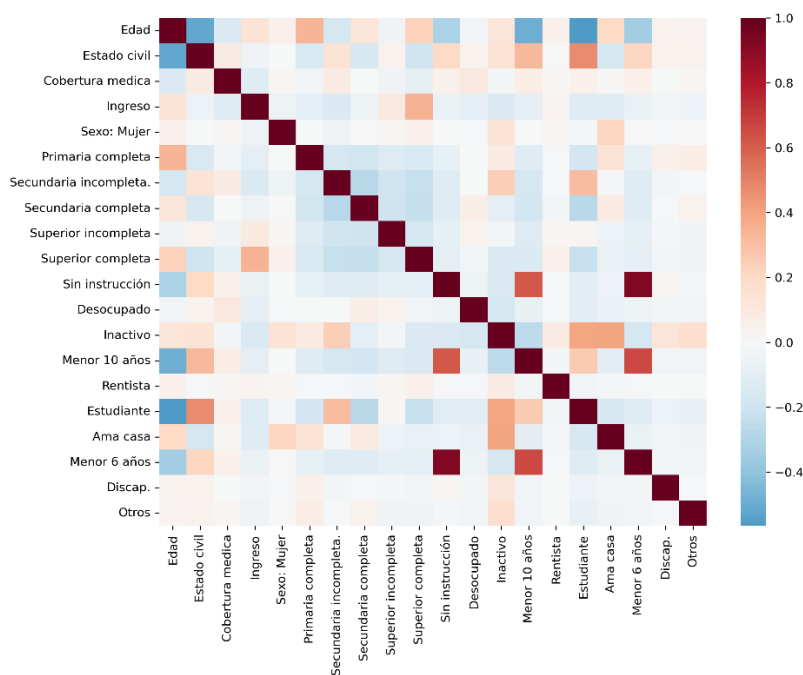
- El **nivel educativo superior completado** mantiene una correlación positiva moderada con el **ingreso per cápita familiar (IPCF)** en ambos años, más fuerte en 2025.
- Las correlaciones entre **sexo** y el resto de las variables son bajas, indicando que no existen diferencias lineales marcadas por género en las condiciones promedio.
- En ambos años, ser mujer y ama de casa mantienen una correlación positiva moderada, pero más débil en 2025.

Gráfico 2. Matriz de correlación de variables sociodemográficas – Gran Buenos Aires (2005)



Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

Gráfico 3. Matriz de correlación de variables sociodemográficas – Gran Buenos Aires (2025)



Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

3.1 Medición de pobreza y estructura de ingresos

En esta sección avanzamos hacia la identificación de los hogares pobres en el Gran Buenos Aires, tomando como base la información de ingresos totales familiares (ITF) y los valores de la Canasta Básica Total (CBT) para los años 2005 y 2025.

Primero, analizamos la calidad de respuesta de la encuesta en torno a las variables económicas. Encontramos que 21 personas no respondieron su condición de actividad, y que 2.985 observaciones (**17,9 % de la muestra total**) registraron ITF = 0, es decir, no declararon sus ingresos familiares, mientras que **13.680** observaciones sí lo hicieron. A partir de esta distinción, generamos dos bases:

- respondieron, que contiene las observaciones con ITF > 0, y
- norespondieron, que agrupa los hogares con ITF igual a cero.

Estas bases se exportaron para su posterior análisis comparativo y control de calidad.

Posteriormente, incorporamos a la base la variable adulta_equiv, utilizando la tabla de equivalencias oficial (tabla_adulto_equiv.xlsx), que asigna valores diferenciados según sexo y

edad. Por ejemplo, un varón de 2 años equivale a 0,46 adultos y una mujer de 70 años a 0,67. Luego, sumamos los valores individuales de cada hogar mediante groupby sobre las variables CODUSU y NRO_HOGAR, obteniendo el total de adultos equivalentes por hogar en una nueva columna denominada ad_equiv_hogar.

Con esta información, calculamos el ingreso mínimo necesario que cada hogar debía alcanzar para no ser considerado pobre. Multiplicamos el total de adultos equivalentes por el valor de la Canasta Básica Total (CBT) correspondiente a cada año:

- **205,07** pesos en el primer trimestre de 2005, y
- **365.177** pesos en el primer trimestre de 2025.

El resultado de esta operación se almacenó en la variable ingreso_necesario, que representa el umbral de pobreza de cada hogar.

Finalmente, creamos la variable dicotómica pobre, que toma el valor 1 si el ingreso familiar total (ITF) del hogar es menor al ingreso necesario, y 0 en caso contrario. Esto nos permitió estimar la cantidad y el porcentaje de personas en situación de pobreza para cada año:

Tabla 2. Proporción de pobres por año – Gran Buenos Aires (2005 y 2025)

Año	Cantidad de pobres	Total, observaciones	% pobres
2005	2506	9484	26.42%
2025	4204	7181	58.54%

Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

Estos resultados muestran un aumento del porcentaje de pobreza en la región, pasando de un **26.42 %** en 2005 a **58.54 %** en 2025, lo que sugiere un deterioro relativo en la capacidad adquisitiva de los hogares, incluso a pesar del incremento nominal de los ingresos. La mayor proporción de hogares con ingresos no declarados también refleja las dificultades de la EPH para captar adecuadamente la distribución del ingreso en los sectores de menor poder adquisitivo.

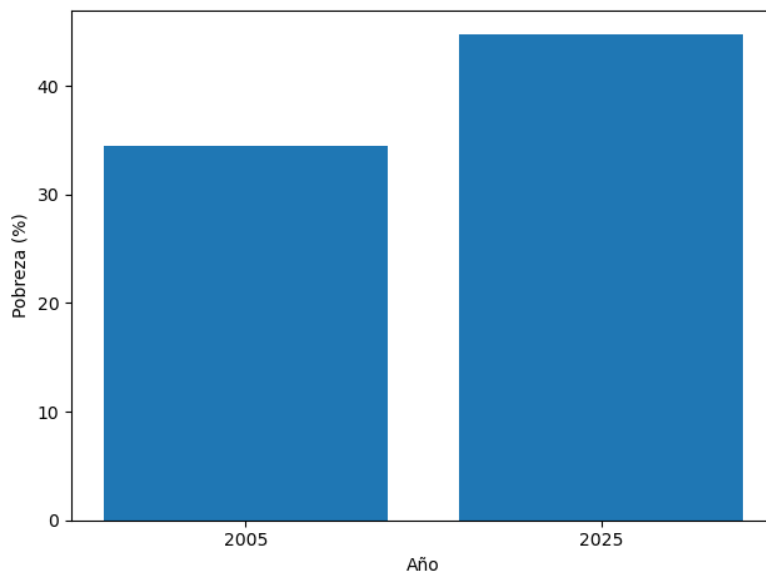
3.2 Análisis descriptivo de la pobreza

Finalmente realizamos un análisis descriptivo de la pobreza incorporando los factores de expansión muestral provistos por la EPH. Esta ponderación nos permitió extrapolar los resultados a la población total del Gran Buenos Aires, garantizando una medición más representativa de la situación socioeconómica.

En primer lugar, estimamos la tasa de pobreza ponderada para cada año. Los resultados muestran que en 2005 la pobreza alcanzaba aproximadamente el **34,5 %** de la población, mientras que en 2025 esta proporción ascendió al **44,7 %**. Este incremento de más de diez

puntos porcentuales refleja un deterioro en las condiciones de vida de los hogares, aun cuando los ingresos nominales hayan aumentado en términos absolutos.

Gráfico 4. Tasa de pobreza por año – Gran Buenos Aires (2025)

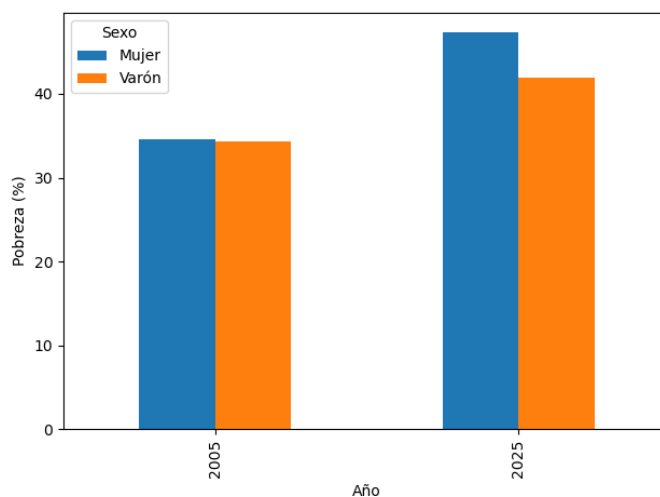


Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

El gráfico 4 evidencia este incremento entre ambos períodos, mostrando un aumento sostenido de la incidencia de la pobreza en el Gran Buenos Aires. Este resultado es coherente con los hallazgos previos sobre la evolución del ingreso familiar y las correlaciones socioeconómicas observadas en las secciones anteriores.

A continuación, analizamos la distribución de la pobreza según el sexo de las personas. Los resultados ponderados indican que en 2005 la pobreza afectaba de manera similar a varones y mujeres, pero en 2025 se amplía la brecha. En ese año, la tasa de pobreza fue de alrededor del **41,9 %** entre los varones y del **47,3 %** entre las mujeres.

Gráfico 5. Tasa de pobreza por año y sexo – Gran Buenos Aires (2025)



Fuente: Elaboración propia a partir de la EPH 2005 Y 2025.

En conjunto, los resultados ponderados confirman una tendencia al aumento de la pobreza y evidencian que los hogares con jefatura femenina se encuentran en una situación más vulnerable. Estos hallazgos refuerzan la importancia de considerar la composición del hogar, el nivel educativo y la situación ocupacional como factores clave en el análisis multidimensional de la pobreza.