# Exploración de Brechas en Ingresos Per Cápita Familiar: Datos de la EPH

#### Introducción

Este artículo presenta un estudio que utiliza datos de la EPH para examinar la brecha de ingresos per cápita familiar promedio entre hombres y mujeres en diferentes regiones del país. A través de un enfoque metodológico que combina el uso de SQL y Excel, se realiza una limpieza y análisis de los datos, con el objetivo de identificar patrones significativos y generar conclusiones relevantes sobre la desigualdad de ingresos. Este análisis no solo permite profundizar en la comprensión de las dinámicas económicas de Argentina, sino que también resalta la importancia de utilizar bases de datos sólidas para abordar cuestiones sociales críticas.

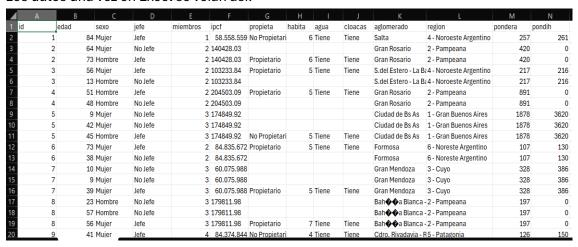
# Limpieza de datos

El proceso de análisis comenzó con la obtención de la base de datos en formato .dta, la cual se convirtió a un archivo CSV para su posterior manejo en Excel. Esta fase inicial requirió un meticuloso trabajo de limpieza de datos, donde abordé desafíos como la identificación de registros faltantes, la eliminación de espacios en blanco y la conversión de columnas en formatos adecuados, siendo la columna de ingreso per cápita familiar (IPCF) la más crítica. Para asegurar la integridad de los datos, se realizó una transformación de los valores de IPCF a formato numérico, y se creó una nueva columna donde dividí el IPCF original por mil. Este procedimiento garantizó que los datos estuvieran en condiciones óptimas para su análisis posterior.

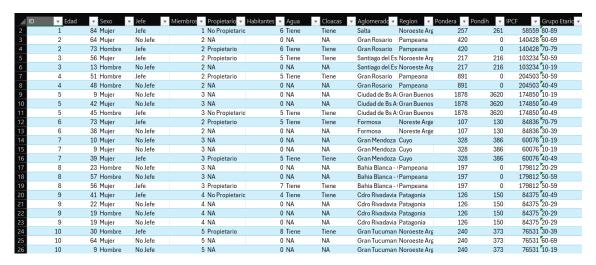
#### Sintaxis en STATA:

use "C:\Users\Franc\Downloads\arg\_2020\_ephc\_t3.dta", clear export delimited "C:\Users\Franc\Downloads\arg\_2020\_ephc\_t3.csv", replace}

#### Los datos una vez en Excel se veían así:



Una vez procesados, los datos se veían así posteriormente:



## Consultas en SQL Server

Con los datos en orden, los exporté a SQL Server para realizar un análisis más eficiente. Allí, implementé indicadores básicos, tales como:

- Total de observaciones de la base de datos.
- Edad promedio.
- IPCF promedio.
- Porcentaje de hombres y mujeres.
- IPCF promedio segmentado por género y grupo etario.

Además, añadí una nueva columna denominada "grupo etario" para facilitar la segmentación de los datos. Posteriormente, utilicé esta información para crear gráficos y tablas dinámicas en Excel.

La sintaxis utilizada fue la siguiente:

```
■USE eph
 SELECT * FROM eph data;
 -- numero de obersaciones en la EPH (personas) no expandido
 SELECT COUNT(*) FROM eph data;
-- respuesta: 97684
 -- Edad Promedio
 SELECT AVG(edad) FROM eph_data;
-- respuesta: 36
 -- Porcentaje de Mujeres y Hombres
≐SELECT
     COUNT(CASE WHEN sexo = 'Mujer' THEN 1 END)*100/COUNT(*) as porcentaje_mujeres,
     COUNT(CASE WHEN sexo = 'Hombre' THEN 1 END)*100/COUNT(*) AS porcentaje_hombres
 FROM eph_data;
⊡-- respuesta_ 52% de mujeres y 47% de hombres
 -- porcentaje de hombres y mujeres para los diferentes grupos etarios
 -- Para el total de Argentina
```

```
-- porcentaje de hombres y mujeres para los diferentes grupos etarios
 -- Para el total de Argentina
SELECT
    CASE
         WHEN edad BETWEEN 0 AND 9 THEN '[0-9]'
         WHEN edad BETWEEN 10 AND 19 THEN '[10-19]'
         WHEN edad BETWEEN 20 AND 29 THEN '[20-29]'
         WHEN edad BETWEEN 30 AND 39 THEN '[30-39]'
         WHEN edad BETWEEN 40 AND 49 THEN '[40-49]'
         WHEN edad BETWEEN 50 AND 59 THEN '[50-59]'
         WHEN edad BETWEEN 60 AND 69 THEN '[60-69]'
         WHEN edad BETWEEN 70 AND 79 THEN '[70-79]'
         WHEN edad BETWEEN 80 AND 89 THEN '[80-89]'
         ELSE '[90 y mas]'
     END AS Grupo_Etario,
         (COUNT(CASE WHEN sexo = 'Mujer' THEN 1 END)*100/COUNT(*)) as Porcentaje_Mujeres,
         (COUNT(CASE WHEN sexo = 'Hombre' THEN 1 END)*100/COUNT(*)) AS Porcentaje_Hombres
FROM
    eph_data
GROUP BY
         WHEN edad BETWEEN 0 AND 9 THEN '[0-9]'
         WHEN edad BETWEEN 10 AND 19 THEN '[10-19]'
         WHEN edad BETWEEN 20 AND 29 THEN '[20-29]'
         WHEN edad BETWEEN 30 AND 39 THEN '[30-39]'
         WHEN edad BETWEEN 40 AND 49 THEN '[40-49]'
         WHEN edad BETWEEN 50 AND 59 THEN '[50-59]' WHEN edad BETWEEN 60 AND 69 THEN '[60-69]'
         WHEN edad BETWEEN 70 AND 79 THEN '[70-79]'
         WHEN edad BETWEEN 80 AND 89 THEN '[80-89]'
         ELSE '[90 y mas]'
    END
ORDER BY
    Grupo_Etario;
```

## Análisis

Lo primero fue trabajar con la tabla de los grupos etarios que exporte desde SQL Server hacia Excel para poder realizar un grafico ilustrativo y sacar conclusiones.

Etiquetas de fila	Promedio de IPCF
0-9	\$ 61.973,38
10-19	\$ 65.866,17
20-29	\$ 82.964,52
30-39	\$ 95.014,04
40-49	\$ 98.289,86
50-59	\$ 116.556,56
60-69	\$ 130.982,13
70-79	\$ 131.157,98
80-89	\$ 142.196,51
90 y más	\$ 147.518,69
Total general	\$ 94.488,13

La tabla muestra el Ingreso Per Cápita Familiar (IPCF) promedio en pesos argentinos, segmentado por grupos etarios. A partir de esta información, se pueden realizar varias observaciones interesantes:

A medida que aumenta el grupo etario, los ingresos per cápita familiares también crecen. Por ejemplo, los niños de 0-9 años tienen un IPCF de \$61.973,38, mientras que las personas de 90 años o más alcanzan un ingreso promedio de \$147.518,69.

Los mayores aumentos en los ingresos se observan a partir del grupo de 50-59 años (\$116.556,56), superando el promedio general de \$94.488,13.

Este crecimiento continúa hasta el grupo etario de 90 años o más, lo que sugiere que los ingresos familiares tienden a aumentar significativamente en personas mayores, posiblemente debido a factores como jubilaciones, ahorro acumulado, o la salida de hijos del hogar, dejando más ingresos disponibles por persona.

Los grupos más jóvenes (0-29 años) muestran los ingresos per cápita más bajos, con valores entre \$61.973,38 y \$82.964,52. Esto podría estar relacionado con familias más numerosas (más hijos, menor ingreso per cápita) o con el hecho de que los padres en estos grupos etarios todavía están en las primeras etapas de sus carreras laborales.

Las personas en el rango de 90 años o más tienen el ingreso promedio más alto (\$147.518,69), lo que puede estar relacionado con familias más pequeñas o con el hecho de que, en muchos casos, los adultos mayores reciben ingresos de fuentes como pensiones o renta.

El IPCF promedio en toda la población es de \$94.488,13, lo que indica que, en general, la mayor parte de los grupos etarios por encima de los 30 años están por encima de este promedio, mientras que los más jóvenes (0-29 años) están por debajo.



# Brechas de género en el ingreso per cápita familiar por región

Un aspecto que consideré muy importante fue analizar si existía una brecha significativa en los ingresos per cápita familiar promedio entre hombres y mujeres, tanto en general como por región. Para ello, generé una tabla dinámica que nos permite obtener observaciones relevantes para el análisis. (pag )

#### Descripción de los resultados:

#### Cuyo:

Los ingresos son similares entre hombres (\$84.974,42) y mujeres (\$82.818,93), aunque los hombres ganan en promedio un poco más.

Gran Mendoza muestra la menor brecha entre géneros dentro de esta región.

#### **Gran Buenos Aires:**

Aquí la diferencia es más notoria: los hombres ganan \$118.099,62 en promedio, mientras que las mujeres ganan \$111.905,41. La diferencia es mayor en la Ciudad de Buenos Aires (\$173.033,14 vs \$162.536,66), lo que implica que el núcleo urbano más grande del país muestra una significativa disparidad.

#### Noreste Argentino:

Esta región muestra una menor diferencia entre hombres y mujeres, siendo Corrientes y Formosa ejemplos de paridad. Sin embargo, Gran Resistencia muestra una notable diferencia, con los hombres ganando considerablemente más que las mujeres.

#### Noroeste Argentino:

Aquí las brechas son también significativas, aunque algunas áreas como Gran Tucumán - T. Viejo presentan una leve ventaja a favor de las mujeres. En general, esta región presenta una diferencia menor en comparación con otras regiones.

#### Pampeana:

El promedio general es relativamente equilibrado (\$98.152,94 para hombres y \$96.513,92 para mujeres), aunque Bahía Blanca - Cerri muestra un ingreso mayor para las mujeres. Sin embargo, otras áreas como Gran La Plata y Gran Rosario mantienen una brecha en favor de los hombres.

#### Patagonia:

Esta región presenta la mayor brecha entre hombres y mujeres, con los hombres ganando en promedio \$124.316,81 y las mujeres \$119.632,69. Ushuaia - Río Grande y Comodoro Rivadavia muestran grandes diferencias, con ingresos particularmente altos para los hombres.

Pro	medio	de l	IPCF

Etiquetas de fila	Hombre	Mujer	Total general
Cuyo	\$ 84.974,42	\$ 82.818,93	\$ 83.836,91
Gran Mendoza	\$91.040,21	\$ 88.122,47	\$89.464,52
Gran San Juan	\$ 76.167,36	\$ 75.216,02	\$ 75.668,69
San Luis - El Chorrillo	\$ 87.316,32	\$ 84.283,83	\$85.760,75
Gran Buenos Aires	\$ 118.099,62	\$ 111.905,41	\$ 114.871,85
Ciudad de Bs As	\$ 173.033,14	\$ 162.536,66	\$ 167.439,46
Partidos del GBA	\$ 102.905,24	\$ 97.021,31	\$99.859,01
Noreste Argentino	\$80.601,60	\$ 78.749,31	\$ 79.624,56
Corrientes	\$80.862,01	\$80.415,79	\$80.622,84
Formosa	\$ 81.822,57	\$80.002,08	\$80.864,24
Gran Resistencia	\$ 68.712,63	\$ 63.527,48	\$ 65.949,96
Posadas	\$ 91.505,37	\$ 92.857,69	\$ 92.202,07
Noroeste Argentino	\$ 73.396,69	\$ 71.850,46	\$ 72.598,81
Gran Catamarca	\$ 76.108,31	\$71.790,71	\$ 73.904,10
Gran Tucuman - T Viejo	\$ 71.168,65	\$ 72.338,64	\$71.786,31
Jujuy – Palpala	\$ 79.383,55	\$77.496,80	\$ 78.447,55
La Rioja	\$ 65.463,95	\$62.381,00	\$ 63.872,85
Salta	\$80.960,81	\$ 79.253,22	\$80.084,93
Santiago del Estero - La Banda2	\$ 64.231,23	\$65.034,72	\$ 64.657,44
Pampeana	\$ 98.152,94	\$ 96.513,92	\$ 97.300,86
Bahia Blanca – Cerri	\$ 110.949,17	\$ 114.850,58	\$ 112.993,58
Concordia	\$ 67.995,01	\$68.612,80	\$ 68.317,24
Gran Cordoba	\$ 91.804,95	\$89.642,01	\$ 90.665,57
Gran La Plata	\$ 103.349,92	\$ 96.806,15	\$ 99.973,27
Gran Parana	\$ 101.470,87	\$ 101.493,21	\$ 101.482,87
Gran Rosario	\$ 108.055,29	\$ 105.769,46	\$ 106.850,36
Gran Santa Fe	\$ 92.544,46	\$89.917,20	\$ 91.171,83
Mar del Plata – Batan	\$ 100.276,63	\$ 100.478,71	\$ 100.379,45
San Nicolas villa Constitucion	\$ 92.979,35	\$ 90.805,26	\$ 91.889,42
Santa Rosa – Toay	\$ 121.838,65	\$ 116.141,67	\$ 118.938,10
Rio Cuarto	\$ 102.674,51	\$ 102.782,15	\$ 102.730,06
Patagonia	\$ 124.316,81	\$ 119.632,69	\$ 121.912,37
Comodoro Rivadavia - Rada Tilly	\$ 136.748,34	\$ 126.883,04	\$ 131.657,83
Neuquen Plottier	\$ 120.904,11	\$ 119.135,63	\$ 120.004,03
Rawson Trelew	\$ 109.475,22	\$ 107.526,51	\$ 108.483,95
Viedma Carmen de Patagones	\$ 102.008,93	\$ 106.334,08	\$ 104.277,83
Ushuaia - Rio Grande	\$ 156.431,19	\$ 142.470,59	\$ 149.217,68
Rio Gallegos	\$ 129.166,92	\$ 120.635,85	\$ 124.879,92
Fotal general	\$ 95.968,73	\$ 93.121,84	\$ 94.488,13

#### Conclusiones:

En cuanto a la brecha de género, la diferencia en los ingresos entre hombres y mujeres es consistente a nivel nacional, pero varía significativamente entre regiones. La Patagonia es la región con las brechas más amplias, especialmente en ciudades como Ushuaia y Comodoro Rivadavia, donde los hombres ganan mucho más que las mujeres. Esto podría estar relacionado con sectores económicos dominados por hombres, como la industria petrolera y minera. Algunas regiones, como el Noreste Argentino y ciertas áreas de la región Pampeana, muestran una mayor equidad en ingresos entre hombres y mujeres, lo cual podría ser un indicio de menores desigualdades de género o mayor participación femenina en sectores laborales mejor remunerados. Aunque el ingreso promedio es más alto en la Ciudad de Buenos Aires, la brecha de género también es significativa. Este dato podría reflejar la concentración de empleos bien remunerados en sectores donde predominan los hombres.

# Acceso: Agua y Cloacas por Aglomerado

En base a la base de datos me interese por buscar si existían diferencias en la relación entre el acceso a servicios públicos (agua y cloacas) y los ingresos per cápita familiares promedio en diferentes aglomerados urbanos de Argentina.

Aglomerado	Porcentaje con Agua	Porcentaje con Cloacas	ipcf Promedio
Ciudad de Bs As	44,5%	44,1%	\$ 31.694,2
Bahia Blanca – Cerri	41,6%	34,9%	\$ 35.981,6
Santa Rosa – Toay	39,9%	35,5%	\$ 37.201,5
Ushuaia - Rio Grande	39,8%	38,0%	\$ 35.968,5
Rio Gallegos	38,7%	37,9%	\$ 36.103,0
Rio Cuarto	38,5%	33,9%	\$ 36.738,8
Gran Rosario	38,2%	30,9%	\$ 38.433,1
Rawson Trelew	38,0%	35,7%	\$ 35.490,5
Gran Parana	37,6%	32,7%	\$ 36.941,2
Mar del Plata – Batan	37,4%	33,9%	\$ 37.867,2
Viedma Carmen de Patagones	37,2%	36,8%	\$ 37.399,7
Neuquen Plottier	37,2%	35,3%	\$ 35.060,5
Gran La Plata	35,8%	23,9%	\$ 37.554,8
Posadas	35,8%	18,0%	\$ 38.541,7
Gran Cordoba	35,6%	15,5%	\$ 37.197,9
Gran Santa Fe	35,3%	19,8%	\$ 39.317,8
San Luis - El Chorrillo	34,6%	33,6%	\$ 39.645,2
San Nicolas Villa Constitucion	34,1%	29,6%	\$ 38.867,6
Cdro. Rivadavia - R.Tilly	34,0%	32,1%	\$ 33.438,8
Gran Mendoza	33,8%	31,3%	\$ 39.425,1
Partidos del GBA	33,4%	21,0%	\$ 37.493,4
Concordia	33,2%	27,3%	\$ 38.774,1
Gran Resistencia	32,3%	19,1%	\$ 39.917,5
Formosa	31,9%	31,0%	\$ 39.635,2
Corrientes	31,6%	29,7%	\$ 38.985,4
Jujuy – Palpala	31,2%	29,5%	\$ 40.471,4
La Rioja	31,1%	29,0%	\$ 42.876,2
Gran Tucuman - T. Viejo	30,7%	22,2%	\$ 39.644,0
Gran San Juan	30,2%	14,3%	\$ 42.033,1
Gran Catamarca	29,8%	22,6%	\$ 41.579,9
Salta	29,5%	26,6%	\$ 39.618,8
Santiago Del Estero - La Banda	28,6%	15,0%	\$ 41.242,6

#### Conclusiones:

Los aglomerados con mejores porcentajes de acceso a agua y cloacas no necesariamente tienen los IPCF más altos. Por ejemplo, Ciudad de Buenos Aires tiene un 44,5% de acceso a agua y un 44,1% de acceso a cloacas, pero su IPCF es de \$31.694,2, que está por debajo del promedio de otras ciudades como La Rioja (\$42.876,2) con un acceso a agua del 31,1% y cloacas del 29,0%.

Hay un claro desequilibrio en el acceso a servicios esenciales como agua y cloacas, con algunas ciudades como Posadas que solo tienen un 18% de acceso a cloacas, pero con un IPCF relativamente alto de \$38.541,7.

Otras ciudades, como Gran Córdoba, muestran un bajo acceso a cloacas (15,5%) pero tienen un ingreso per cápita promedio de \$37.197,9, lo cual revela que el acceso a estos servicios no es necesariamente un indicador directo del nivel de ingreso.

Los aglomerados con menor acceso tanto a agua como cloacas, como Santiago del Estero - La Banda (28,6% acceso a agua, 15% acceso a cloacas), tienen un IPCF de \$41.242,6, lo que sugiere que, a pesar de tener menos infraestructura, pueden tener un ingreso promedio relativamente alto comparado con otras ciudades.

Regiones urbanas más grandes como el Gran Buenos Aires (GBA) presentan porcentajes bajos en acceso a servicios básicos (33,4% acceso a agua y 21% a cloacas), mientras que tienen un IPCF promedio de \$37.493,4, lo que implica que incluso en áreas metropolitanas importantes, el acceso a infraestructura básica puede estar rezagado en relación con el nivel de ingresos.

Esta disparidad sugiere la necesidad de políticas públicas que mejoren el acceso a agua y cloacas, especialmente en aquellas áreas que, a pesar de tener ingresos relativamente altos, carecen de estos servicios esenciales. El hecho de que varias ciudades con IPCF altos tengan un bajo acceso a infraestructura podría ser un factor que limite su potencial de desarrollo a largo plazo.

# Acceso: Agua y Cloacas por Región

No solo se realizaron observaciones para los diferentes aglomerados, sino que también se tuvo en cuenta el porcentaje con acceso al agua y cloacas por Región.

Región	de Porcentaje con agua	
Cuyo		0,33
Gran Buenos Aires		0,36
Noreste Argentino		0,33
Noroeste Argentino		0,30
Pampea		0,37
Patagonia		0,37
Total general		2,06

La **Patagonia** es la región con el mayor porcentaje acumulado de acceso a agua, con un valor de 0,3733 (37,33%), lo que sugiere una mayor cobertura de acceso a agua en esta región en comparación con otras.

La **región Pampeana** también presenta un valor alto de 0,3674 (36,74%), lo que indica una buena cobertura de agua.

*Gran Buenos Aires* sigue con un 35,9%, lo que es notable considerando la densidad poblacional de esta área.

#### Regiones con Menor Acceso:

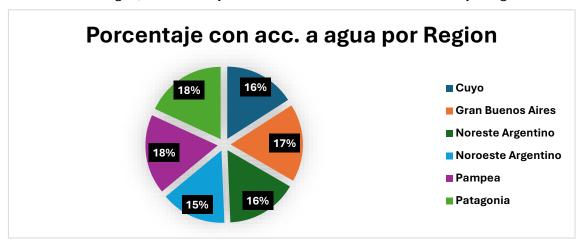
**Noroeste Argentino** tiene el menor porcentaje de acceso a agua con 0,3014 (30,14%), lo que refleja una posible disparidad en el acceso a servicios básicos como el agua en esta región.

**Noreste Argentino** tiene un porcentaje ligeramente mayor, con 0,3291 (32,91%), pero sigue estando por debajo del promedio nacional.

La **suma total** de los porcentajes es 2,0579, lo que sugiere que, si calculamos el promedio por región, este sería:

**Promedio total** = 2,05796 / 6 = 0,343 (34,3%).

Esto indica que, en promedio, alrededor del 34,3% de la población en estas regiones tiene acceso al agua, lo cual refleja una cobertura considerablemente baja en general.



Porcentaje – Acceso a cloacas por región.

	Porcentaje con Acceso a	
Región	Cloacas	
Cuyo		0,26
Gran Buenos Aires		0,26
Noreste Argentino		0,24
Noroeste Argentino		0,24
Pampea		0,28
Patagonia		0,36
Total general		1,64

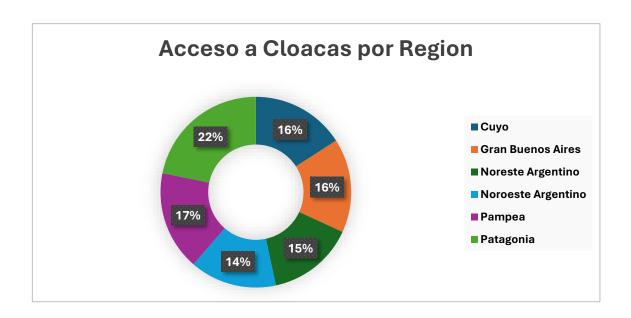
La **Patagonia** es la región con el mayor porcentaje de acceso a cloacas, con un valor de 0,36 (36%), lo que sugiere que esta región tiene la mejor cobertura en comparación con las demás.

La **región Pampeana** también muestra un porcentaje considerable, alcanzando 0,28 (28%), lo que indica una cobertura razonable, aunque aún podría mejorarse.

**Gran Buenos Aires y Cuyo** presentan un acceso similar, ambos con 0,26 (26%). Esto es notable, dado que Gran Buenos Aires es una de las áreas más densamente pobladas del país, lo que puede presentar desafíos en la implementación de servicios básicos.

**Noroeste Argentino y Noreste Argentino** tienen los porcentajes más bajos de acceso a cloacas, ambos con 0,24 (24%). Esto refleja una disparidad significativa en el acceso a servicios de saneamiento básico en estas regiones, lo que podría tener implicaciones en la salud pública.

La **suma total** de los porcentajes es 1,64 (164%), lo que, al dividirlo por el número de regiones, da un promedio total de 0,273 (27,3%). Esto indica que, en promedio, aproximadamente el 27,3% de la población en estas regiones tiene acceso a cloacas, lo que refleja una cobertura bastante baja en general.



#### Conclusiones:

El análisis realizado sobre los datos de la Encuesta Permanente de Hogares (EPH) ha permitido obtener conclusiones significativas respecto a los ingresos per cápita familiar promedio en Argentina. A lo largo de este proceso, comenzamos con una cuidadosa limpieza de datos en Excel, donde se abordaron problemas como registros faltantes y la conversión de la columna de ingreso per cápita familiar (IPCF) a un formato numérico adecuado. Esta etapa fue crucial, ya que aseguraba que los datos estuvieran listos para un análisis riguroso.

Una vez completada la limpieza, los datos fueron importados a SQL Server, donde realizamos cálculos de indicadores clave. Estos incluyeron la edad promedio y el IPCF promedio desglosados por grupo etario, así como un análisis detallado de las diferencias en ingresos entre hombres y mujeres. Al segmentar estos ingresos por aglomerado y región, se evidenció que la brecha de ingresos es notoria, especialmente en ciertos grupos etarios, lo que sugiere la necesidad de políticas que aborden estas disparidades.

Además, se generaron tablas dinámicas que facilitaron la visualización de los datos, permitiendo observar patrones y tendencias que contribuyen a un mejor entendimiento de la situación socioeconómica en el país. Este enfoque analítico no solo iluminó las diferencias en los ingresos promedio per cápita familiar, sino que también destacó la importancia de considerar factores regionales y de género en el estudio de la desigualdad económica.