

```
In [17]: import pandas as pd
import numpy as np
import os
```

```
In [5]: import pandas as pd

poblacion = pd.read_csv('C:/Users/Lautaro Perez/Documents/Teclab/Análisis_y_Visu
esperanza = pd.read_csv('C:/Users/Lautaro Perez/Documents/Teclab/Análisis_y_Visu
hogares = pd.read_csv('C:/Users/Lautaro Perez/Documents/Teclab/Análisis_y_Visual
```

```
In [6]: poblacion.head()
```

```
Out[6]:
```

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres
0	Total País	2010	40788453	19940704	20847749
1	Total País	2011	41261490	20180791	21080699
2	Total País	2012	41733271	20420391	21312880
3	Total País	2013	42202935	20659037	21543898
4	Total País	2014	42669500	20896203	21773297

```
In [7]: poblacion.tail()
```

```
Out[7]:
```

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres
770	Tierra del Fuego	2036	241593	122567	119026
771	Tierra del Fuego	2037	245734	124625	121109
772	Tierra del Fuego	2038	249853	126670	123183
773	Tierra del Fuego	2039	253948	128702	125246
774	Tierra del Fuego	2040	258020	130721	127299

```
In [8]: poblacion.shape
```

```
Out[8]: (775, 5)
```

```
In [9]: poblacion.dtypes
```

```
Out[9]: provincia      object
anio                  int64
poblacion_total      int64
poblacion_varones    int64
poblacion_mujeres    int64
dtype: object
```

```
In [10]: poblacion.describe()
```

Out[10]:

	anio	poblacion_total	poblacion_varones	poblacion_mujeres
count	775.000000	7.750000e+02	7.750000e+02	7.750000e+02
mean	2025.000000	3.777746e+06	1.856888e+06	1.920858e+06
std	8.950048	9.560571e+06	4.699604e+06	4.861043e+06
min	2010.000000	1.316610e+05	6.723500e+04	6.442600e+04
25%	2017.000000	5.845510e+05	2.906740e+05	2.934905e+05
50%	2025.000000	1.017731e+06	5.061010e+05	5.161370e+05
75%	2033.000000	1.855285e+06	9.138865e+05	9.404745e+05
max	2040.000000	5.277848e+07	2.603809e+07	2.674038e+07

In [11]:

```
pobla_filtrado = poblacion[poblacion['provincia'] != 'Total País']
pobla_filtrado
```

Out[11]:

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres
31	Capital Federal	2010	3028481	1405566	1622915
32	Capital Federal	2011	3033639	1409835	1623804
33	Capital Federal	2012	3038860	1414105	1624755
34	Capital Federal	2013	3044076	1418339	1625737
35	Capital Federal	2014	3049229	1422507	1626722
...
770	Tierra del Fuego	2036	241593	122567	119026
771	Tierra del Fuego	2037	245734	124625	121109
772	Tierra del Fuego	2038	249853	126670	123183
773	Tierra del Fuego	2039	253948	128702	125246
774	Tierra del Fuego	2040	258020	130721	127299

744 rows × 5 columns

In [12]:

```
poblac_superf = pd.merge(pobla_filtrado, hogares, left_on=['provincia'], right_o
poblac_superf
```

Out[12]:

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres	provincia_
0	Capital Federal	2010	3028481	1405566	1622915	
1	Capital Federal	2011	3033639	1409835	1623804	
2	Capital Federal	2012	3038860	1414105	1624755	
3	Capital Federal	2013	3044076	1418339	1625737	
4	Capital Federal	2014	3049229	1422507	1626722	
...	
739	Tierra del Fuego	2036	241593	122567	119026	9
740	Tierra del Fuego	2037	245734	124625	121109	9
741	Tierra del Fuego	2038	249853	126670	123183	9
742	Tierra del Fuego	2039	253948	128702	125246	9
743	Tierra del Fuego	2040	258020	130721	127299	9

744 rows × 10 columns



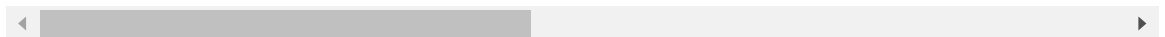
In [13]:

```
poblac_superf['densidad_de_poblacion'] = poblac_superf['poblacion_total'] / pobl  
poblac_superf
```

Out[13]:

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres	provincia_
0	Capital Federal	2010	3028481	1405566	1622915	
1	Capital Federal	2011	3033639	1409835	1623804	
2	Capital Federal	2012	3038860	1414105	1624755	
3	Capital Federal	2013	3044076	1418339	1625737	
4	Capital Federal	2014	3049229	1422507	1626722	
...	
739	Tierra del Fuego	2036	241593	122567	119026	9
740	Tierra del Fuego	2037	245734	124625	121109	9
741	Tierra del Fuego	2038	249853	126670	123183	9
742	Tierra del Fuego	2039	253948	128702	125246	9
743	Tierra del Fuego	2040	258020	130721	127299	9

744 rows × 11 columns



In [14]: `poblac_superf.isnull().sum()`

Out[14]:

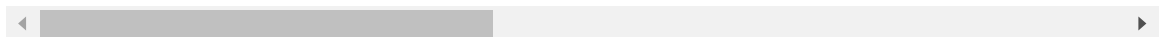
provincia	0
anio	0
poblacion_total	0
poblacion_varones	0
poblacion_mujeres	0
provincia_id	0
hogares	0
viviendas_particulares	0
viviendas_particulares_habitadas	0
superficie_km2	0
densidad_de_poblacion	0
dtype:	int64

In [15]: `poblac_superf['fuera_de_rango'] = (poblac_superf['densidad_de_poblacion'] - np.m
poblac_superf`

Out[15]:

	provincia	anio	poblacion_total	poblacion_varones	poblacion_mujeres	provincia_
0	Capital Federal	2010	3028481	1405566	1622915	
1	Capital Federal	2011	3033639	1409835	1623804	
2	Capital Federal	2012	3038860	1414105	1624755	
3	Capital Federal	2013	3044076	1418339	1625737	
4	Capital Federal	2014	3049229	1422507	1626722	
...	
739	Tierra del Fuego	2036	241593	122567	119026	9
740	Tierra del Fuego	2037	245734	124625	121109	9
741	Tierra del Fuego	2038	249853	126670	123183	9
742	Tierra del Fuego	2039	253948	128702	125246	9
743	Tierra del Fuego	2040	258020	130721	127299	9

744 rows × 12 columns



In [18]: `p99 = np.percentile(poblac_superf['fuera_de_rango'], 99)`
`p99`

Out[18]: `np.float64(4.821306724666531)`

Análisis de la variable "Superficie"

Al analizar la variable superficie km² se evidencia que: El área promedio de las provincias argentinas es de 112,107 km². La distribución de la superficie es asimétrica, con un percentil 25 de 29,801 km², una mediana de 91,302 km² y un percentil 75 de 160,837 km².

Provincias como Capital Federal y Tierra del Fuego tienen áreas pequeñas, mientras que otras son significativamente más extensas, lo que explica la diferencia entre el promedio y la mediana.