




Guia de preguntas - Hogares.csv

1. Análisis de Tendencia Central y Dispersión:

Calcule las medidas de tendencia central (media, mediana) y dispersión (varianza, desviación estándar) para las variables IngresoMensual, GastoMensual, y EdadPromedioHogar. ¿Los valores encontrados son consistentes con lo que esperarías en un contexto económico real?

Name	Distribution	Mean	Mode	Median	Dispersion
 N IngresoMensual		91719.8818	40607.43	91055.67	0.3550
 N GastoMensual		62151.3	20623.9	62352.3	0.401912
 N EdadPromedio...		40.192	36	38	0.411913

Ingreso Mensual: La media y la mediana son muy similares → la distribución es relativamente simétrica.

Sin embargo, la moda es mucho más baja, lo que indica que hay bastantes hogares en niveles de ingresos bajos, pero también existen ingresos altos que elevan el promedio. Esto es consistente con lo que se observa en la economía real: ingresos concentrados en pocos hogares de nivel alto, pero la mayoría en un rango menor.

Gasto Mensual: Media y mediana prácticamente iguales → distribución centrada y estable. La moda mucho más baja muestra que muchos hogares gastan poco, aunque algunos tienen gastos mucho mayores.

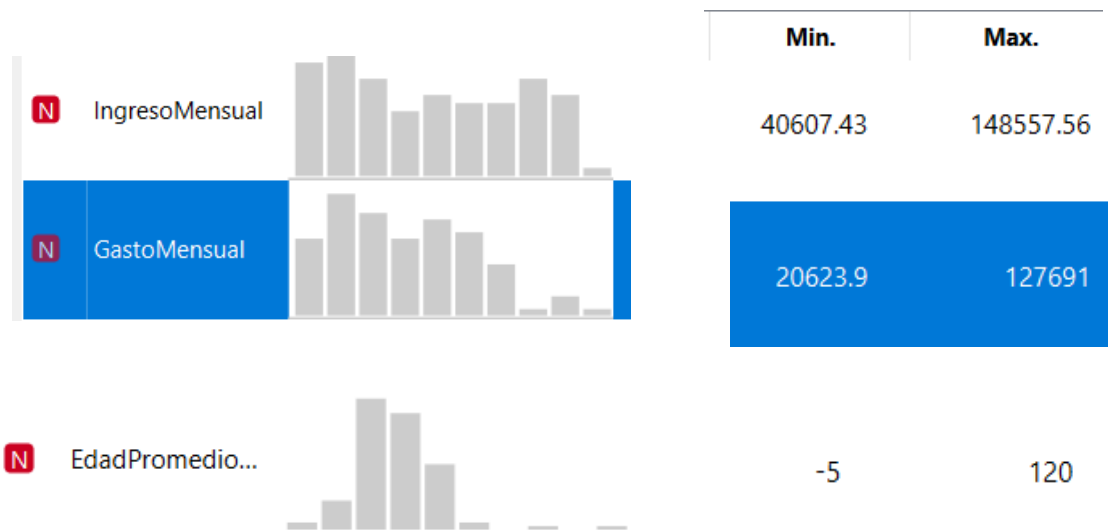
Esto también es razonable: en economía real, el gasto depende del ingreso disponible, y aunque la mayoría gasta en niveles bajos-medios, unos pocos consumen bastante más.

Edad Promedio: La media está cerca de la mediana y la moda → distribución normal centrada en los 36–40 años.

Esto refleja una edad típica para jefes de hogar, consistente con datos demográficos reales.

2. Verificación de Máximos y Mínimos:

Identifique los valores máximos y mínimos para IngresoMensual, GastoMensual, y EdadPromedioHogar. ¿Estos valores son realistas? Discuta cualquier valor que parezca inusual o fuera de lo esperado



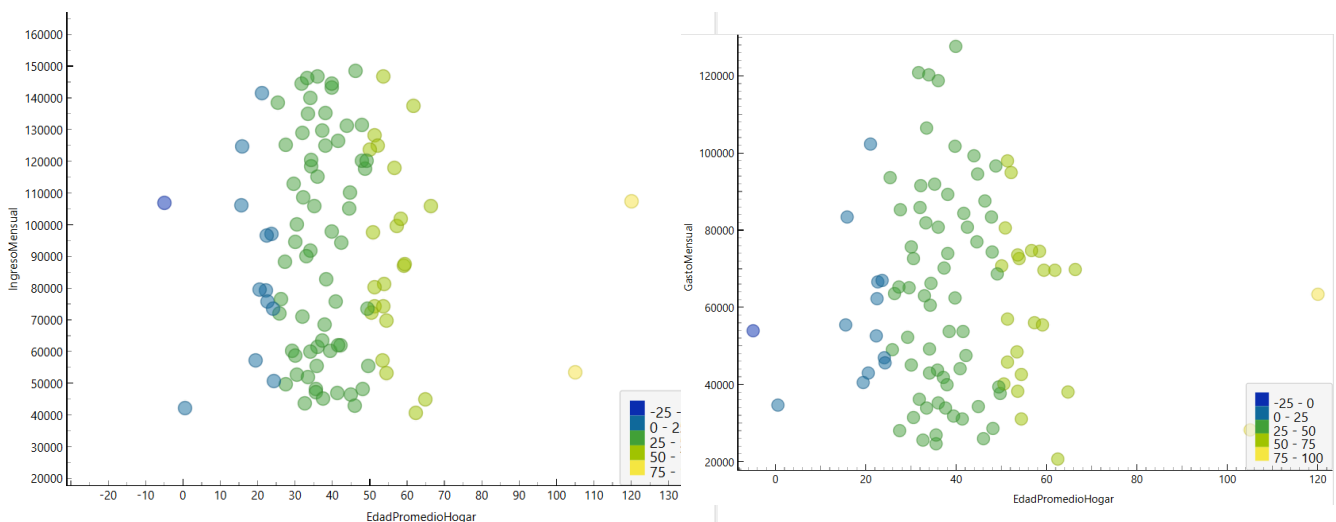
En **IngresoMensual** y **GastoMensual**, los valores resultan **muy bajos para el contexto económico argentino actual**, donde los ingresos y gastos de los hogares suelen ser bastante mayores (ejemplo: salario mínimo es aprox 300.000). Esto sugiere que el dataset puede corresponder a **años anteriores** o estar ajustado de otra forma.

En **EdadPromedioHogar**, el valor **-5** es un error claro, mientras que **120** es un outlier extremo, poco realista.

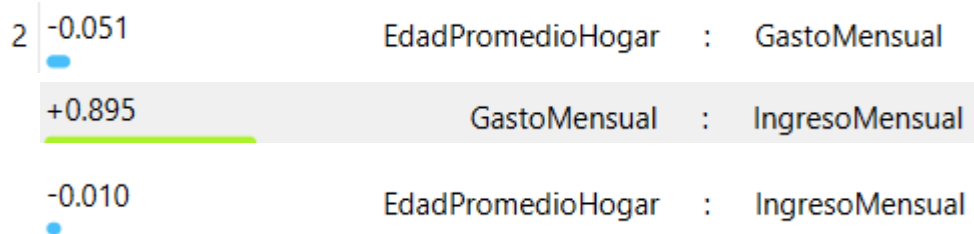
Por lo tanto, aunque los valores siguen una lógica relativa (a mayor ingreso, mayor gasto), en términos absolutos no son consistentes con la situación económica presente.

3. Relación entre Variables:

Explore la relación entre el **EdadPromedioHogar** y otras variables numéricas como **IngresoMensual** y **GastoMensual**. ¿Hay alguna tendencia visible que indique cómo la edad promedio del hogar afecta los ingresos y gastos?

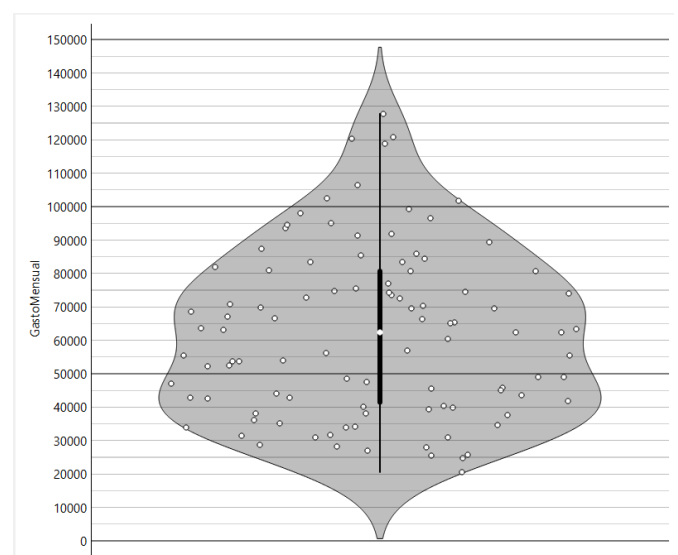
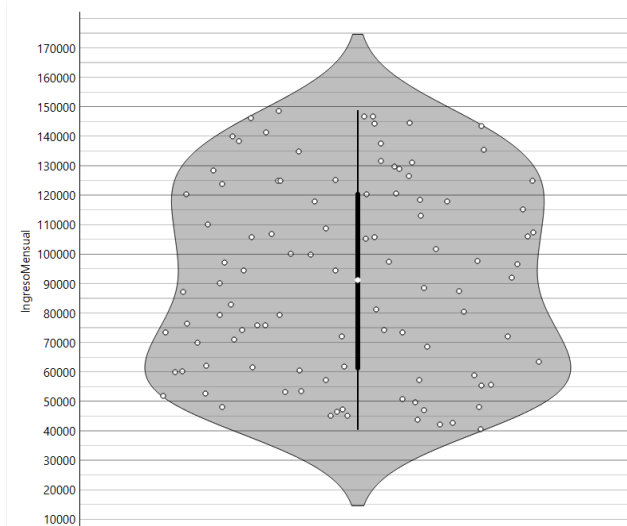


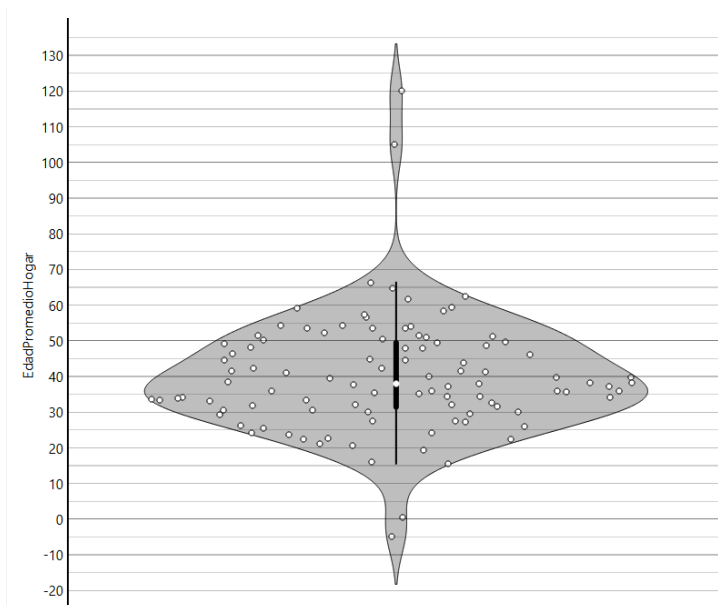
la edad promedio del hogar **no parece ser un predictor fuerte** de los ingresos ni de los gastos, aunque se aprecia que los hogares de edad intermedia concentran los valores más estables, lo cual coincide con lo esperado en la realidad económica: los hogares en edad laboral activa suelen tener los ingresos y gastos más representativos.



El análisis de correlación indica que la EdadPromedioHogar no guarda relación significativa con los ingresos ni con los gastos (coeficientes $-0,010$ y $-0,051$, respectivamente). Esto confirma que la edad no es un factor determinante directo en las variables económicas consideradas. En cambio, se observa una correlación fuerte y positiva entre IngresoMensual y GastoMensual ($+0,895$), lo cual es esperable: a medida que aumentan los ingresos de un hogar, también lo hacen sus niveles de gasto.

4.



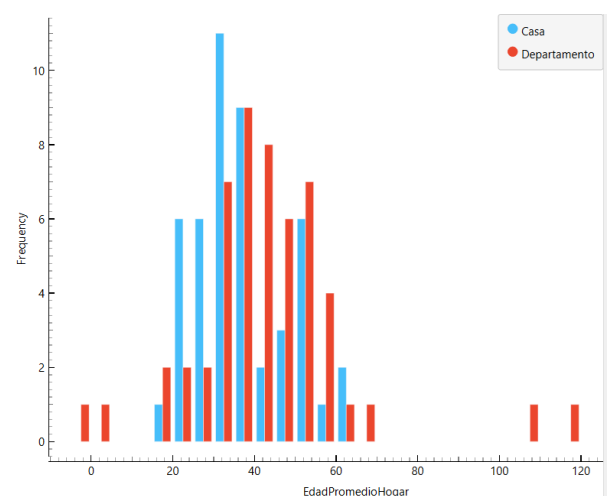
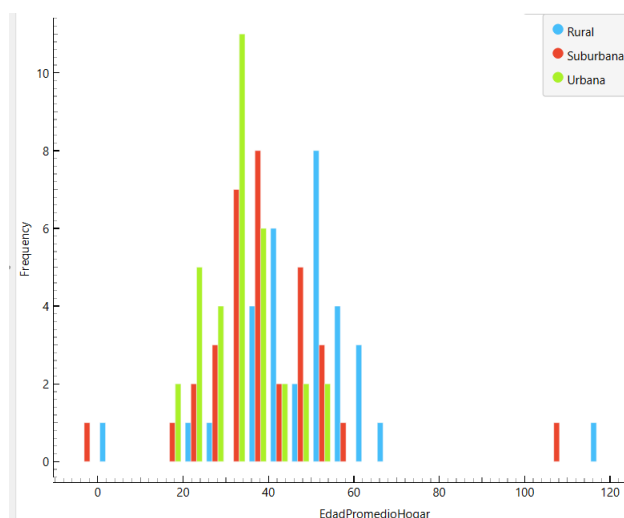


A partir del análisis gráfico se identificaron varios valores atípicos en las variables numéricas. En **IngresoMensual** y **GastoMensual** aparecen hogares con valores extremos (≈ 150.000 y ≈ 127.000) que, aunque posibles en hogares de ingresos altos, se apartan del rango típico de la mayoría y deben ser tratados como outliers.

En **EdadPromedioHogar** se detectaron valores claramente anómalos: un registro negativo (-5), imposible en la realidad, y un valor de 120 años, extremadamente raro y probablemente un error de carga. Estos datos atípicos pueden explicarse por errores de registro.

5. Comparación entre Grupos:

Analice cómo varía la EdadPromedioHogar entre diferentes tipos de vivienda (Casa, Departamento) y zonas de vivienda (Urbana, Suburbana, Rural). ¿Existen diferencias significativas que puedan sugerir patrones demográficos?

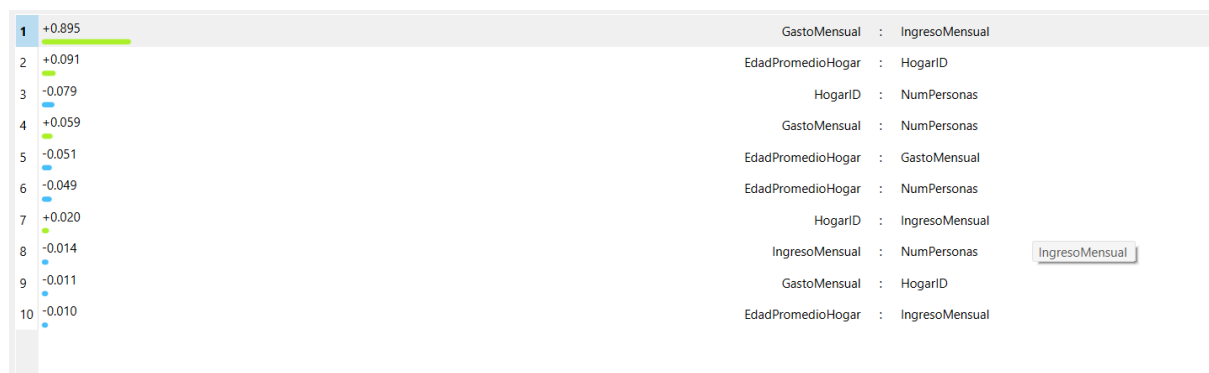


El análisis de la variable EdadPromedioHogar según distintos grupos muestra diferencias demográficas relevantes. En cuanto a la zona de vivienda, los hogares en áreas urbanas tienden a ser más jóvenes, con concentraciones en el rango de 20 a 40 años. Los hogares en áreas suburbanas presentan edades intermedias (30–50 años), mientras que los hogares en zonas rurales muestran promedios más altos (40–60 años) e incluso valores extremos poco frecuentes (mayores de 100 años), lo cual puede deberse a errores de registro o a la presencia de hogares con jefes muy longevos.

Respecto al tipo de vivienda, se observa que los hogares en departamentos corresponden a grupos etarios más jóvenes, mientras que los que habitan en casas presentan edades promedio más elevadas, reflejando hogares familiares más consolidados.

6. Correlaciones:

Investigue la correlación entre las variables numéricas. ¿Qué relaciones se pueden identificar como relevantes y cómo se interpretan?



El análisis de correlaciones evidencia que la relación más importante se da entre **IngresoMensual y GastoMensual (+0,895)**, lo que refleja una correlación positiva muy fuerte y consistente con el comportamiento económico esperado: los hogares con mayores ingresos tienden a gastar más.

En contraste, las demás correlaciones resultan prácticamente nulas. La **EdadPromedioHogar no presenta relación significativa ni con ingresos ni con gastos**, lo que indica que la edad promedio de los integrantes del hogar no es un factor determinante en el nivel económico de este conjunto de datos. Asimismo, el número de personas en el hogar tiene correlaciones muy débiles, apenas sugerentes de que un mayor tamaño del hogar podría asociarse con un leve aumento del gasto.