

## Guía de Preguntas

### 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?

#### IngresoMensual:

Media: 91719,8818

Mediana: 91055,67

Dispersión: 0,3550

#### GastoMensual:

Media: 62151,3

Mediana: 62352,3

Dispersión: 0,401912

#### EdadPromedioHogar:

Media: 40,192

Mediana: 38

Dispersión: 0,411913

Si bien la media y la mediana son valores similares, lo que indica que la distribución es bastante simétrica y sin valores extremos, los datos no resultan consistentes en el contexto argentino 2025. La media de ingresos del hogar (aprox \$91.720) está muy por debajo del Salario Mínimo Vital y Móvil vigente (\$322.200). Esto sugiere que los datos no representan de manera realista la situación económica de los hogares argentinos.

### 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.

IngresoMensual:

Max:148557.56

Min:40607.43

GastoMensual:

Max:129671

Min:20623.9

EdadPromedioHogar:

Max:120

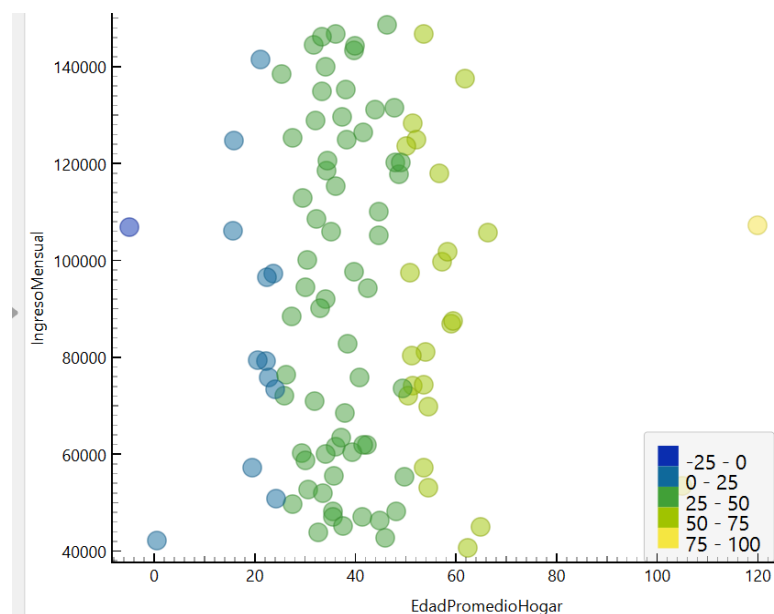
Min:-5

El max y el min de EdadPromedioHogar no son valores realistas. El valor min ( -5) representa una edad negativa, la cual no es posible.

### 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?



### EdadPromedioHogar - IngresoMensual

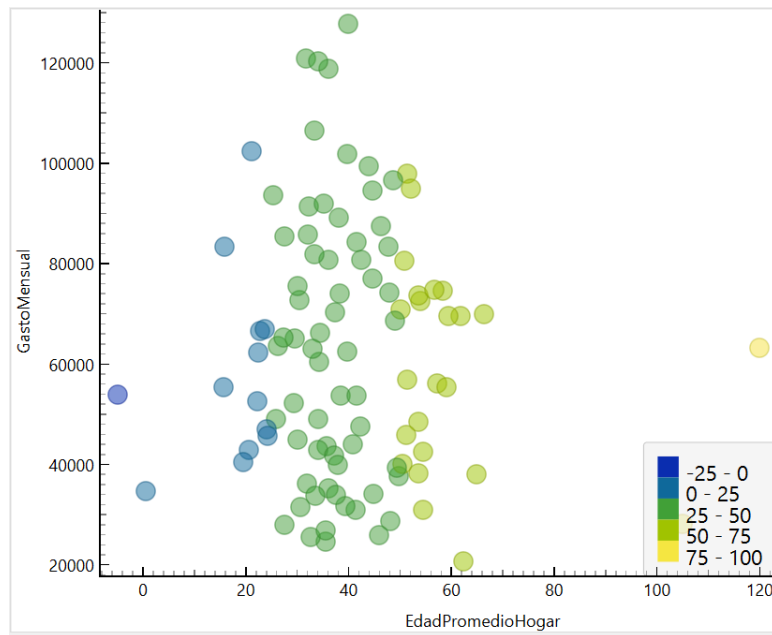
La nube de puntos está bastante dispersa y no se observa una pendiente clara. No hay una relación lineal fuerte: hogares con edades promedio bajas y altas pueden tener ingresos similares.

Se muestran algunos valores atípicos (como ya habíamos mencionado

anteriormente) edad

negativa y la edad

promedio de 120.



### EdadPromedioHogar - GastoMensual

El patrón es similar a **EdadPromedioHogar - IngresoMensual**: no se ve una tendencia clara a que los hogares con mayor edad promedio gasten más o menos.

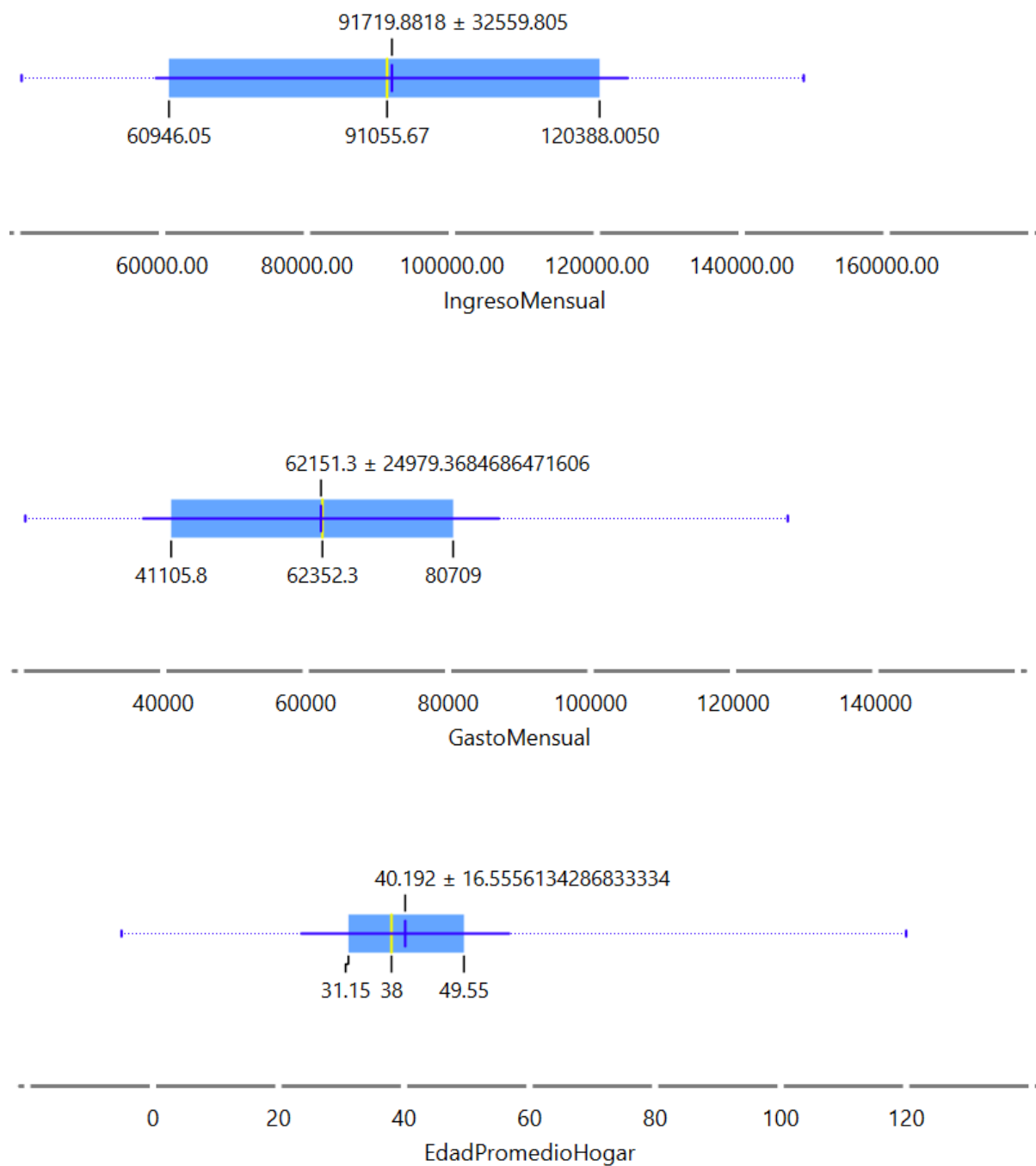
También aparecen los valores mencionados (edad -5 y edad 120).

Por lo analizado en cada gráfico concluimos que no se observa una tendencia visible que indique que la edad promedio del hogar afecte directamente los ingresos ni los gastos.

#### 4. Análisis de Outliers:

Utilizando técnicas gráficas, identificar datos atípicos en las variables numéricas.

Describe por qué estos puntos podrían considerarse atípicos y qué factores podrían explicar estos valores.

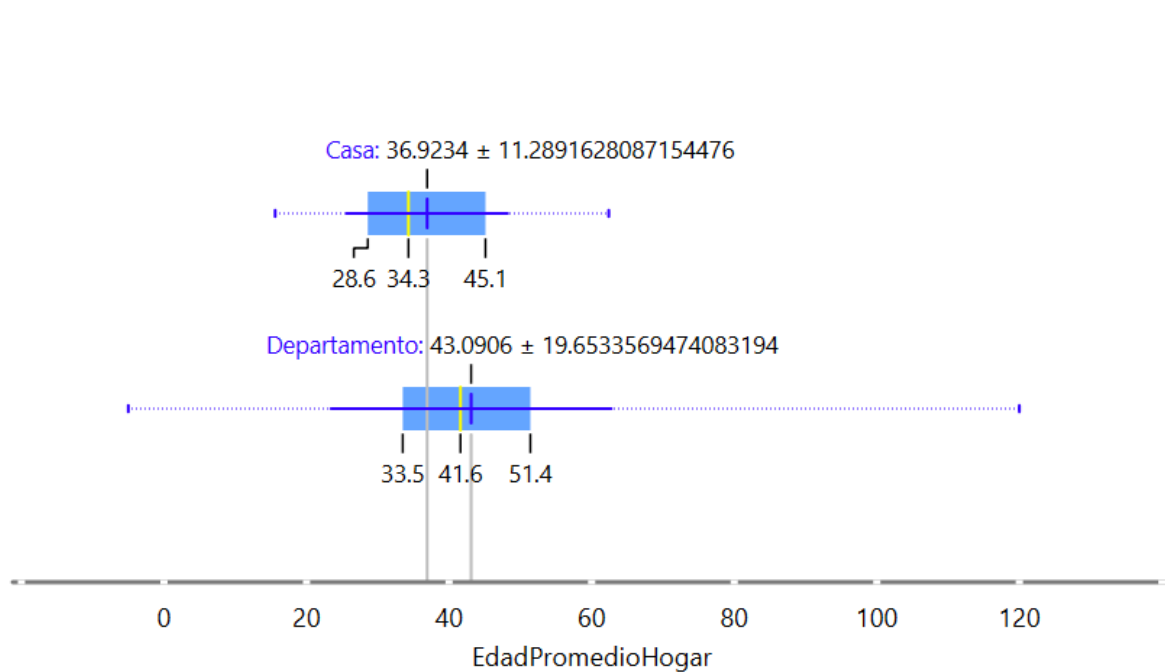


En los gráficos se ven algunos valores que se van muy alejados del resto. En ingresos y gastos aparecen casos con montos mucho más altos que el promedio, que pueden ser familias con mucha más dinero, o personas que gastan más de lo que ganan o simplemente errores de carga.

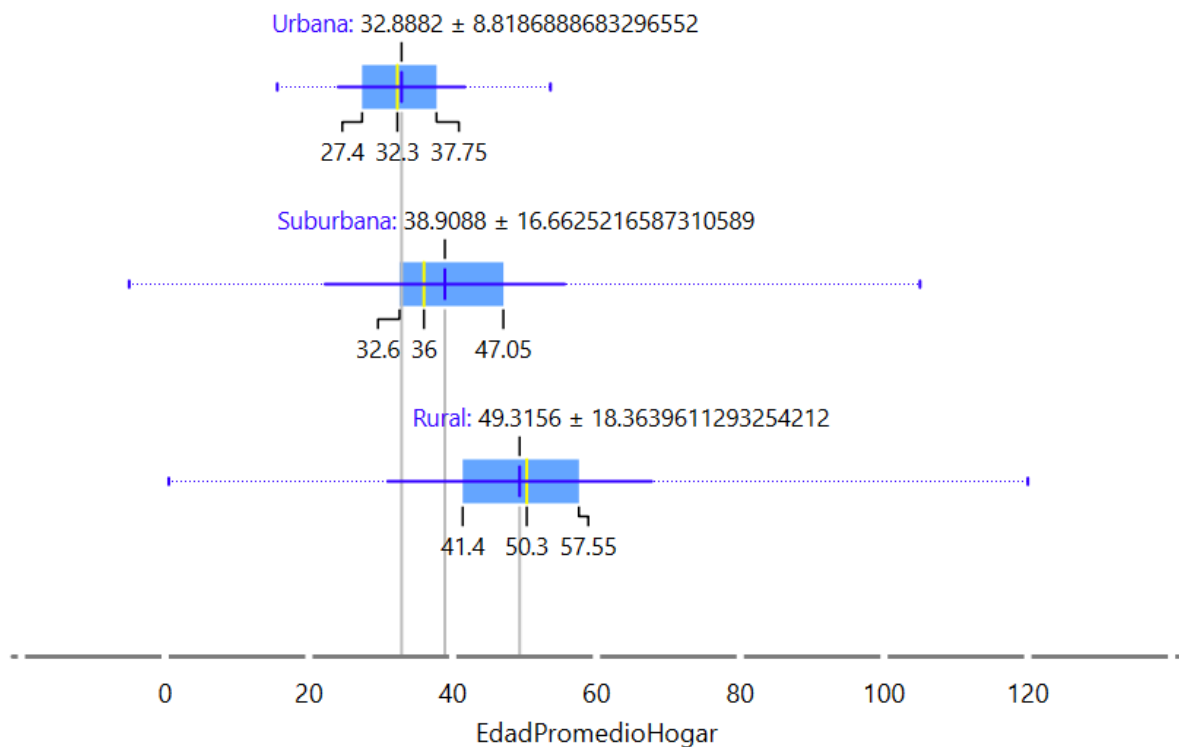
En la edad promedio del hogar hay outliers imposibles, como -5 o 120, que son claramente errores en los datos. Estos valores extremos hacen que el análisis no sea tan confiable y habría que limpiarlos o revisarlos antes de seguir.

### 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?



Student's t: 1.950 (p=0.054, N=100)



ANOVA: 9.612 ( $p=0.000$ ,  $N=100$ )

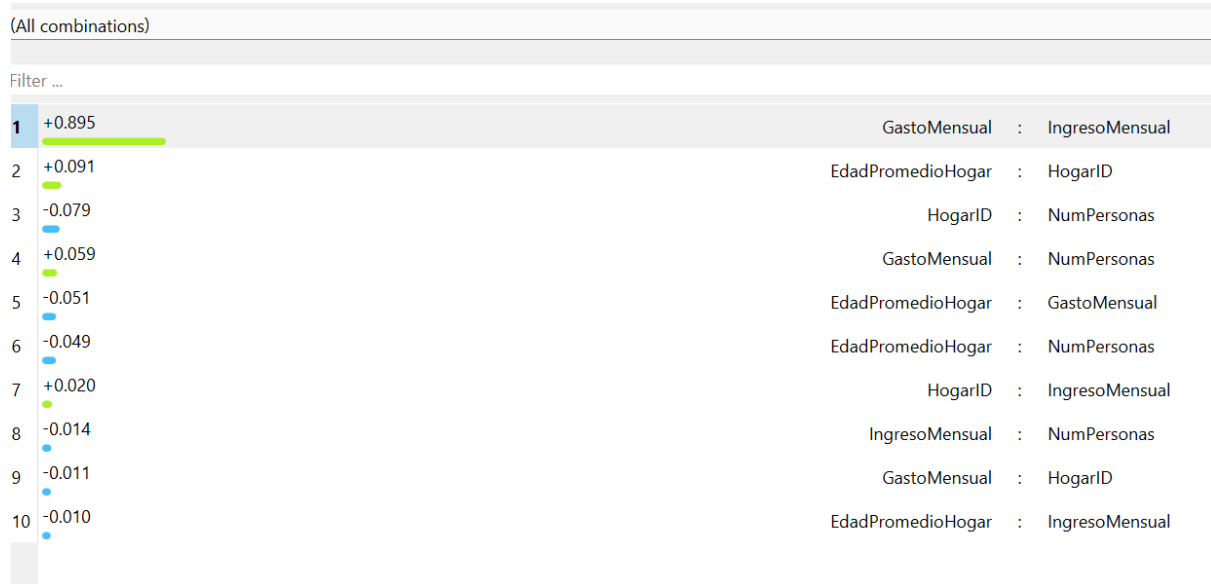
Cuando se compara la edad promedio de los hogares según el tipo de vivienda, se ve que en las casas suelen vivir familias más jóvenes, mientras que en los departamentos la gente es en promedio más grande. Esto tiene sentido porque las casas se encuentran más en zonas suburbanas, donde viven parejas jóvenes con hijos, y los departamentos son más comunes en áreas urbanas con adultos mayores o parejas sin hijos.

Si miramos las zonas, la diferencia es todavía más clara: en las áreas urbanas la edad promedio es más baja ( $\approx 33$  años), en las suburbanas sube un poco ( $\approx 39$  años) y en las rurales es mucho más alta ( $\approx 49$  años). El ANOVA muestra que estas diferencias sí son significativas ( $p=0,000$ ), lo que indica un patrón demográfico fuerte: cuanto más lejos del centro urbano, mayor es la edad promedio de los hogares.

## 6. Correlaciones:

Investigue la correlación entre las variables numéricas.

¿Qué relaciones se pueden identificar como relevantes y cómo se interpretan?



La relación más relevante es la primera correlación de GastoMensual e IngresoMensual con +0,895, que es muy alta. Esto quiere decir que a mayor ingreso, mayor gasto, lo que es algo bastante lógico si lo pensamos en nuestro día a día.

Las demás correlaciones son bastantes bajas, la mayoría muy cercanas a cero. Por lo que no hay relaciones fuertes entre las demás variables analizadas.

Por lo que entendemos que lo más importante que podemos llevarnos del gráfico es que el gasto depende casi directamente del ingreso, el resto no tiene mucha relación.