

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	GastoMensual	EdadPromedioHogar
Media	91719,88	62151,3	40,192
Mediana	91055,67	62352,3	38
Dispersion	0,3550	0,401912	0,411913

Ingresos y gastos son coherentes entre sí (gasto < ingreso, dispersión razonable). En EdadPromedioHogar hay valores fuera de lo común, de ahí la dispersión algo alta.

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	GastoMensual	EdadPromedioHogar
Maximos	148.557,56	127.691	120
Minimos	40.607,63	20.623,9	-5

Los extremos de ingresos y gastos son plausibles. EdadPromedioHogar = -5 y 120 no son realistas.

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?

Con ambas variables la relación es prácticamente nula. Los scatterplots no muestran tendencia clara. La edad promedio del hogar no explica ingresos ni gastos en este set.

4. Análisis de Outliers:

Utilizando técnicas gráficas, identificar datos atípicos en las variables numéricas. Describa por qué estos puntos podrían considerarse atípicos y qué factores podrían explicar estos valores.

Mediante la técnica de box plot, pudimos ver que los atípicos vienen de EdadPromedioHogar (ej.: -5 y 120). Posibles causas: errores de tipeo, o casos ficticios o datos erróneos extremos. Las otras variables no presentan valores atípicos.

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?

Diferentes tipos de vivienda:

Los hogares en departamentos tienen una edad promedio mayor (43,1 años) que los de casas (36,9 años). Además, los departamentos muestran mayor dispersión, indicando una composición etaria más heterogénea que en las casas.

Diferentes zonas de vivienda:

La edad promedio es más baja en zonas urbanas (32,9), sube en suburbanas (38,9) y es más alta en rurales (49,3). La variabilidad también aumenta al alejarse de la ciudad, con las zonas rurales mostrando la mayor dispersión.

6. Correlaciones:

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

La única relación relevante es GastoMensual–IngresoMensual: a mayor ingreso, mayor gasto, con una asociación alta y consistente.

El resto no son relevantes, podemos interpretar entonces que la edad y tamaño del hogar apenas explican ingresos o gastos.