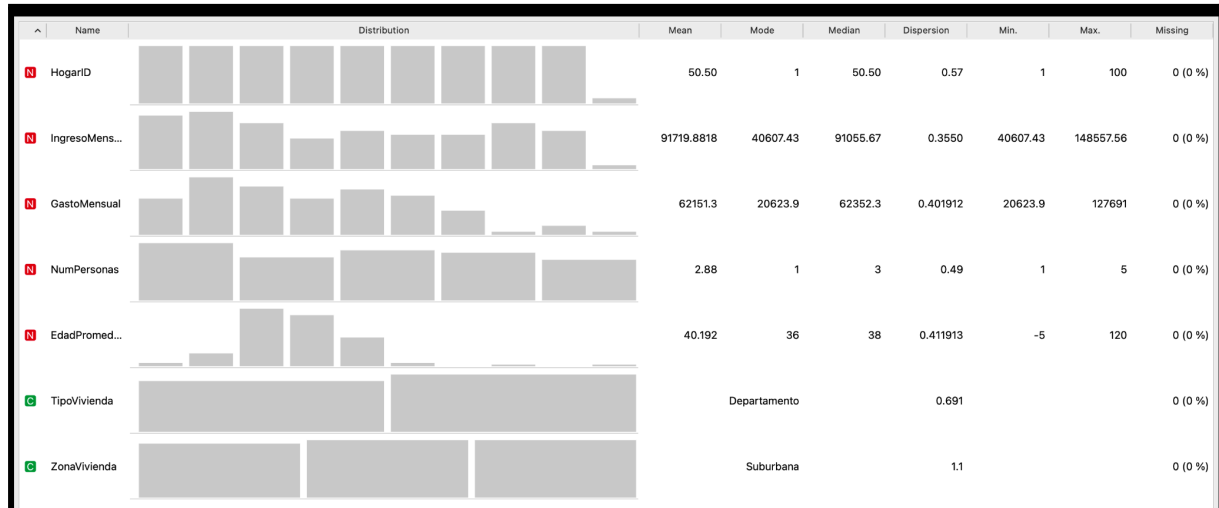


1. Análisis de Tendencia Central y Dispersión:

Usando el widget: Feature Statistic

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?



→ Ingreso mensual

- Media: 91719
- Mediana: 91055
- CV (Dispersion) = 0,3550
- Desvío estándar (σ) $\approx 0,3550 \times 91.719,88 = 32.580$
- Varianza (σ^2) $\approx 32.580^2 = 1.062.916.000$

→ Gasto Mensual

- Media: 62151,3
- Mediana: 62352,3
- Desvío estándar: ≈ 24.960
- Varianza: $\approx 623.000.000$

→ Edad promedio

- Media: 40
- Mediana: 38
- Desvío estándar: $\approx 16,6$ años
- Varianza: ≈ 276

En cuanto al ingreso mensual de los hogares, la media (91.719) y la mediana (91.056) son muy parecidas, lo que indica una distribución equilibrada y sin sesgos.

El desvío estándar, cercano a 32.580, y la varianza, en torno a 1.062 millones, muestran una dispersión considerable, propia de la diferencia económica entre hogares.

Respecto al gasto mensual, se observa una media de 62.151 y una mediana de 62.352, muy cercanas entre sí, lo que refleja estabilidad en la distribución. No obstante, el desvío estándar elevado (≈ 24.960) y la varianza (≈ 623 millones) evidencian diferencias significativas en los niveles de gasto, teniendo sentido con la diversidad socioeconómica de la muestra.

Finalmente, la edad promedio de los hogares presenta una media de 40,2 años y una mediana de 38 años, con un desvío estándar de 16,6 y una varianza de 276. Estas cifras sugieren una composición diversa, abarcando tanto hogares jóvenes como de mayor edad. Los valores de tendencia central se muestran consistentes con lo esperable en un contexto real, mientras que las medidas de dispersión reflejan la variedad propia de una población de hogares.

2. Verificación de Máximos y Mínimos:

Usando el widget: Feature Statistic y Data Table

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.

Max / min

- Ingreso mensual: 40.607 / 148.557
- Gasto mensual: 20.623 / 127.691
- Edad promedio hogar: -5 / 120

En relación a los **ingresos mensuales**, se observa un valor mínimo de 40.607 y un máximo de 148.558. Estos valores resultan posibles dentro de un rango socioeconómico amplio, reflejando desde hogares con recursos más limitados hasta otros con ingresos superiores.

En el caso del **gasto mensual**, los valores extremos se ubican entre 20.624 y 127.691, lo cual se relaciona con la tendencia de que los hogares con mayores ingresos también presentan niveles de gasto más elevados.

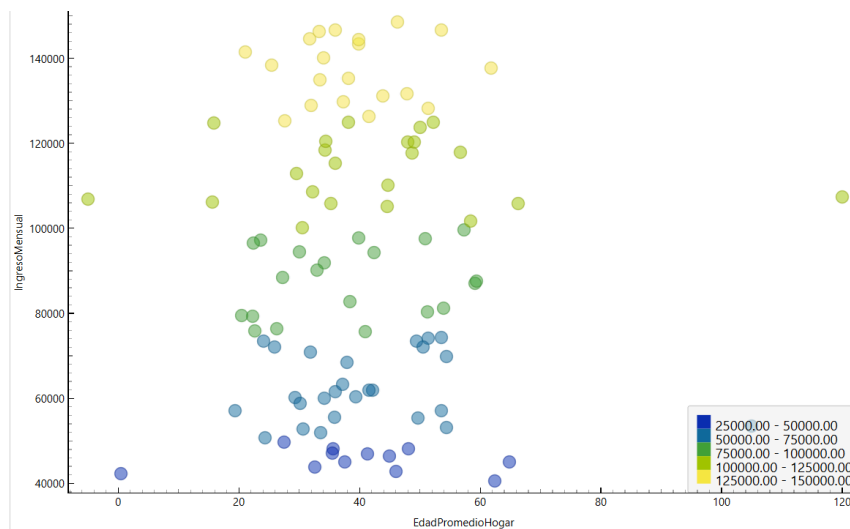
Por último, la **edad promedio** de los hogares presenta un mínimo de -5 años y un máximo de 120 años. Estos valores resultan problemáticos: el mínimo negativo constituye claramente un error de registro, mientras que un promedio de 120 años, aunque es posible, es muy improbable en términos demográficos. Por ello, estos casos deben ser considerados **outliers**.

3. Relación entre Variables

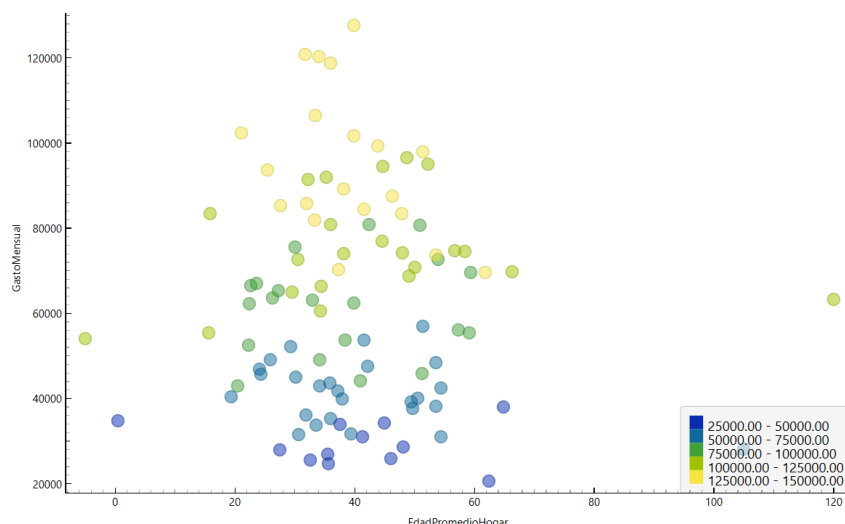
Usando el widget: Scatter Plot y para comprobar un BoxPlot

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 comparado con IngresoMensual



EdadPromedioHogar comparado con GastoMensual

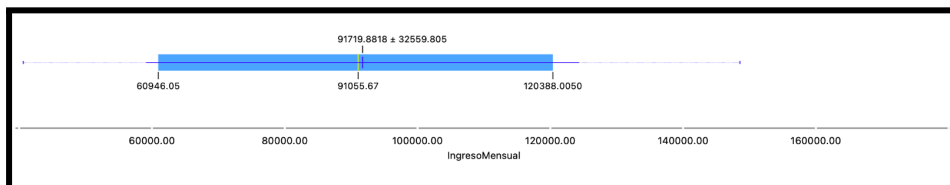
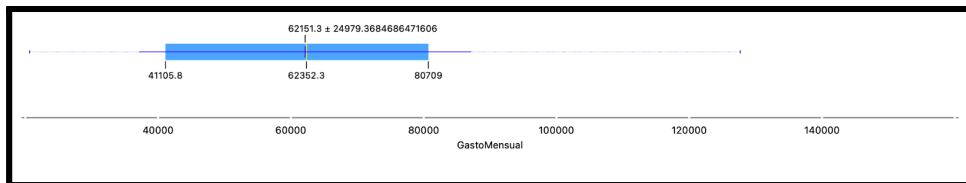
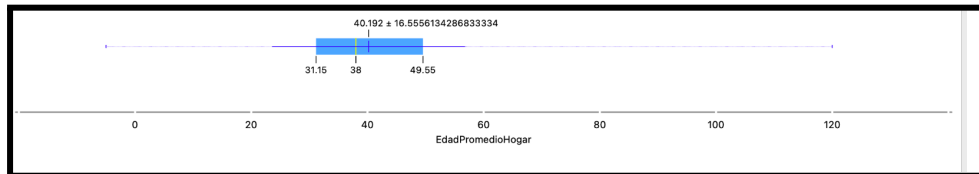


Al analizar la relación entre la edad promedio del hogar y las variables económicas, no se observa un patrón lineal fuerte. En los gráficos de dispersión se ve que los hogares con promedios de edad más bajos (menores de 30 años) suelen tener ingresos y gastos reducidos, lo que resulta lógico si pensamos que en esta etapa los integrantes recién empiezan a insertarse en el mercado laboral. En cambio, en los hogares con edades intermedias (entre 40 y 50 años) aparece más dispersión: algunos presentan ingresos altos y gastos acordes a estos ingresos, mientras que otros se mantienen en niveles más moderados. Finalmente, en los hogares con mayor edad promedio se nota una leve caída tanto en los ingresos como en los gastos, posiblemente asociada al retiro laboral o a una menor actividad económica. En síntesis, la edad promedio puede influir en las dinámicas económicas de los hogares, pero la relación es débil y está atravesada por múltiples factores sociales y contextuales.

4. Análisis de Outliers:

Usando el widget: Box plot, y nos basamos en datos obtenidos del feature statistics

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.

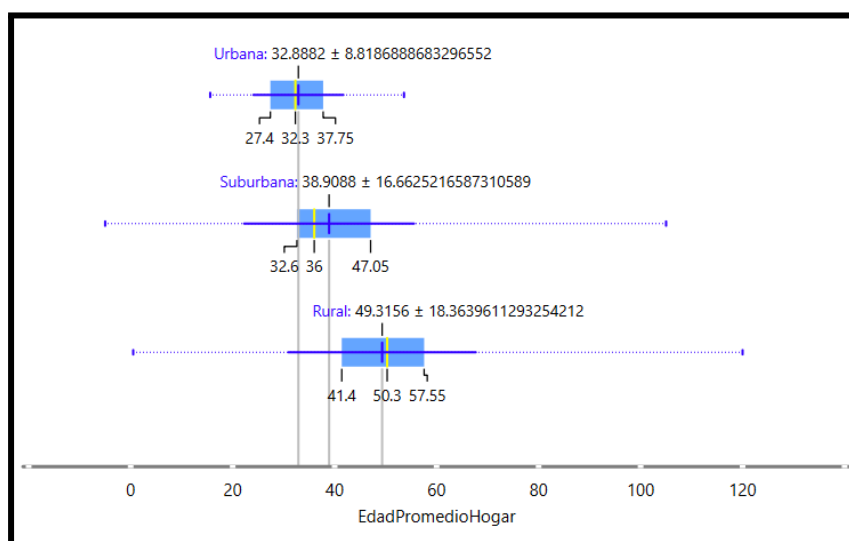
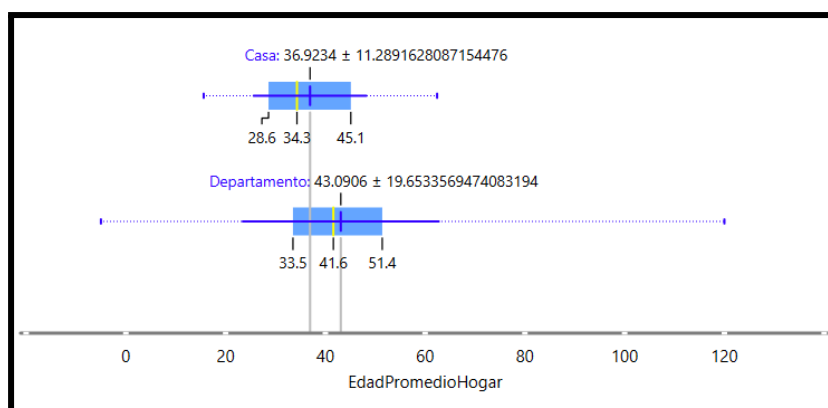


El análisis mediante boxplots muestra un patrón diferenciado según las variables. En la variable **EdadPromedioHogar**, se identifican valores fuera de los límites estadísticos esperados (–5 y 120 años), que constituyen **outliers**. En cambio, tanto en **IngresoMensual** como en **GastoMensual**, los valores extremos se encuentran dentro de rangos aceptables, por lo que no se detectan outliers extremos. Sin embargo, en ambas variables se observa una **alta dispersión**, especialmente en los niveles más altos de ingresos y gastos, lo que refleja la diversidad económica de los hogares sin llegar a ser considerados outliers.

5. Comparación entre Grupos:

Usando el widget: Box plot y distributions para chequear resultados

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** en función del tipo y la zona de vivienda muestran algunas diferencias significativas entre los grupos. En cuanto al **tipo de vivienda**, los hogares que residen en casas presentan una edad promedio menor (media $\approx 36,9$ años), mientras que los que habitan en departamentos, una edad promedio más elevada (media $\approx 43,1$ años). Esta diferencia puede entenderse desde términos demográficos, dado que los departamentos suelen pertenecer a adultos de mayor edad o familias.

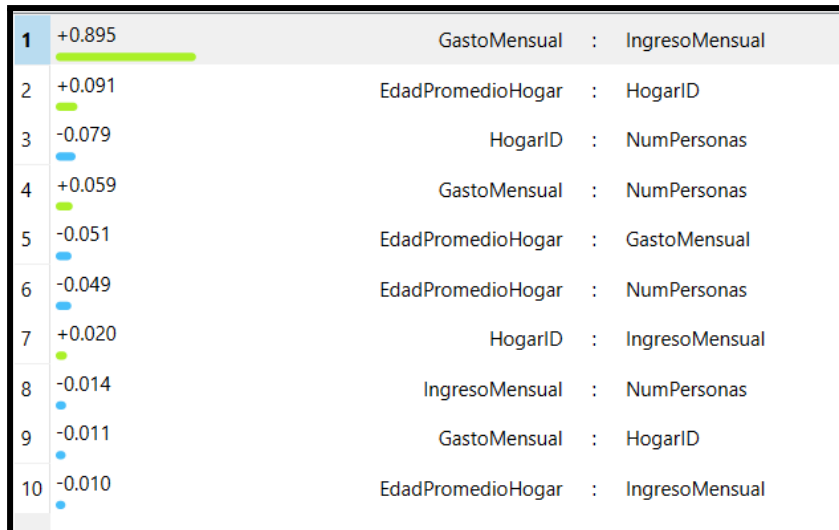
Respecto a la **zona de residencia**, se observan tres perfiles diferenciados: los hogares en **zonas urbanas** tienen una edad promedio más baja ($\approx 32,9$ años), lo que refleja la concentración de población joven en áreas centrales y con mayor movilidad laboral o estudiantil; los hogares en **zonas suburbanas** presentan un promedio intermedio ($\approx 38,9$ años) mientras que los hogares en **zonas rurales** muestran la edad promedio más elevada ($\approx 49,3$ años), probablemente vinculada a la permanencia de adultos mayores en estas áreas y la migración de jóvenes hacia zonas urbanas.

Por lo tanto, con estos resultados podemos ver que tanto el tipo de vivienda como la localización están asociados a patrones demográficos específicos, ayudando a evidenciar la distribución diferencial de los hogares.

6. Correlaciones

Usando el widget: Correlation

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



Al analizar las correlaciones en Orange, lo primero que aparece es la relación entre **IngresoMensual y GastoMensual**, con un coeficiente cercano a 0,90. Se trata de una correlación positiva muy fuerte, lo que significa que los hogares que ganan más también tienden a gastar más. Este resultado es bastante lógico: cuando los ingresos aumentan, también crece la posibilidad de destinar dinero a distintos rubros y, en consecuencia, el nivel de gasto total del hogar.

En cambio, cuando vemos las relaciones de la **EdadPromedioHogar** y del **Número de personas por hogar** con las variables económicas, prácticamente notamos que no hay conexiones relevantes. La edad promedio apenas se relaciona con ingresos y gastos, y el tamaño del hogar tampoco muestra un patrón claro. Esto indica que, dentro de esta muestra, ni la edad promedio ni la cantidad de integrantes alcanzan para explicar diferencias en la situación económica.

En conclusión, lo más importante que muestran las correlaciones es que las variables de **ingreso y gasto van de la mano**, mientras que las demás variables numéricas no parecen tener un peso significativo en estas dinámicas.