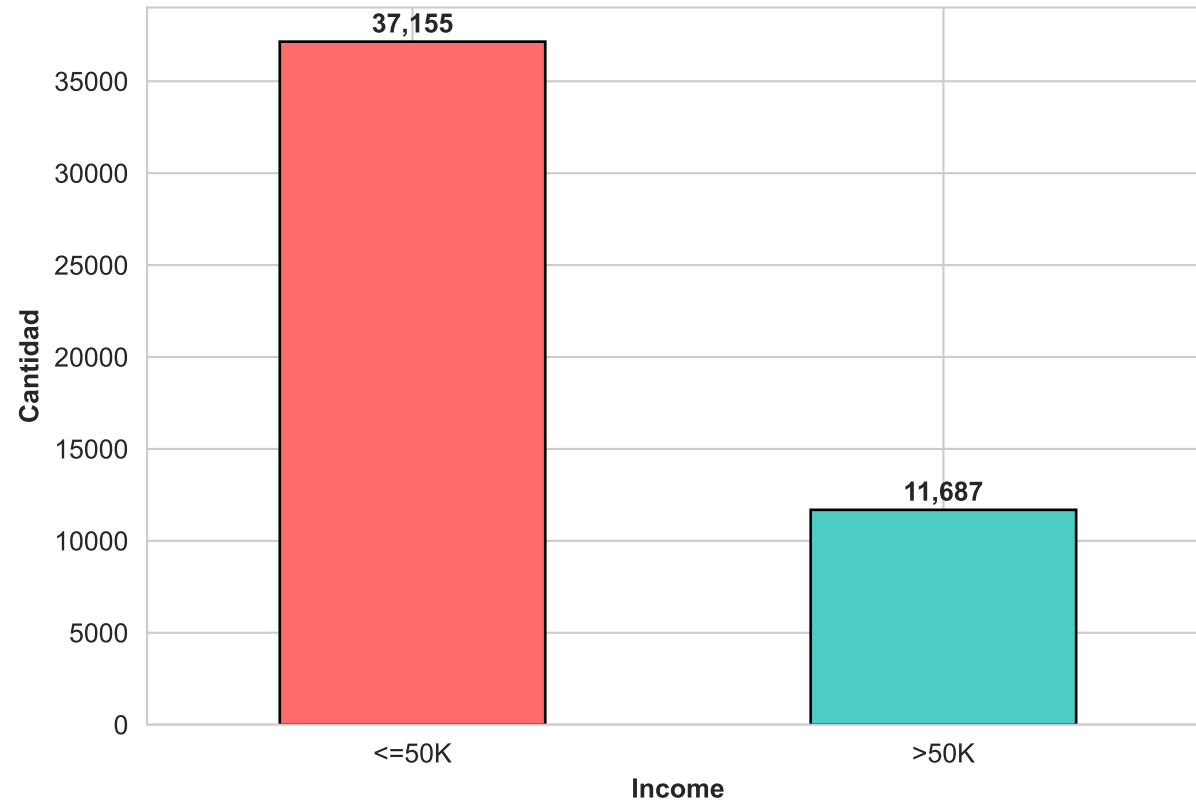
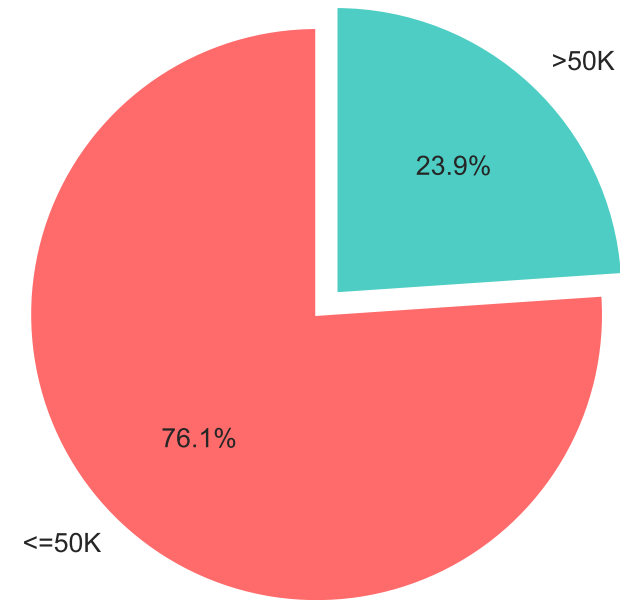


Distribución de la Variable Objetivo: INCOME

Frecuencia Absoluta

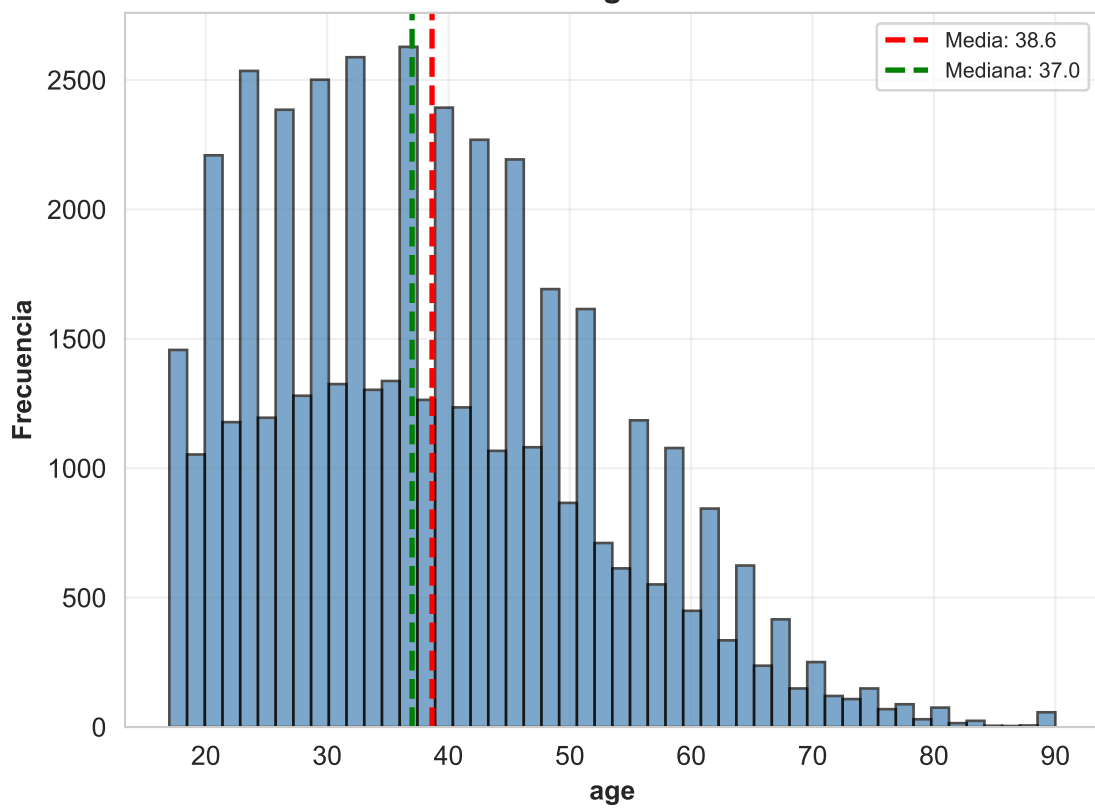


Porcentaje Relativo

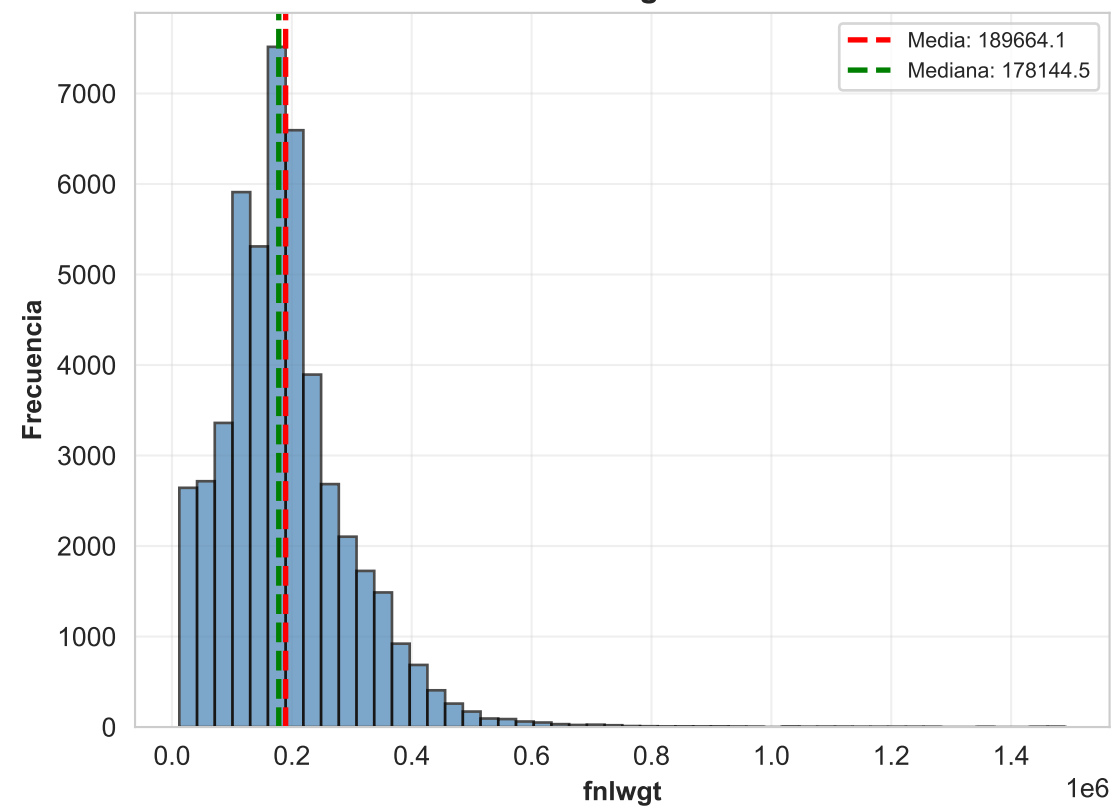


Distribuciones de Variables Numéricas

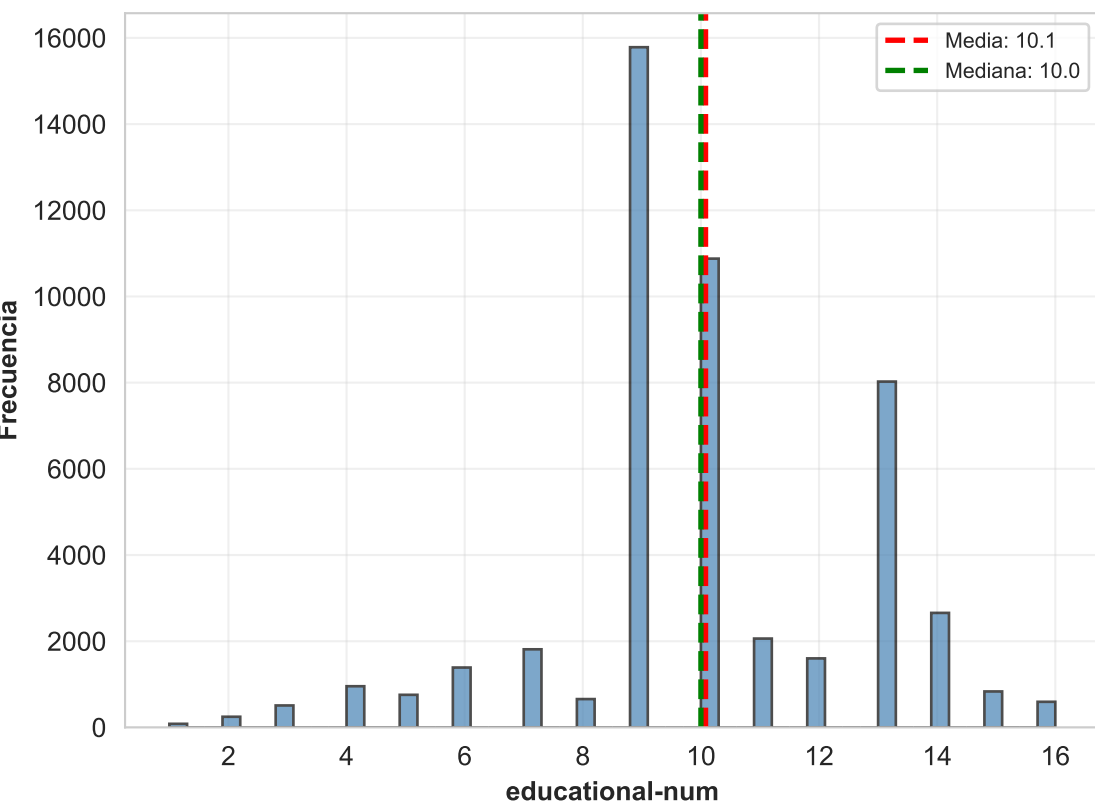
age



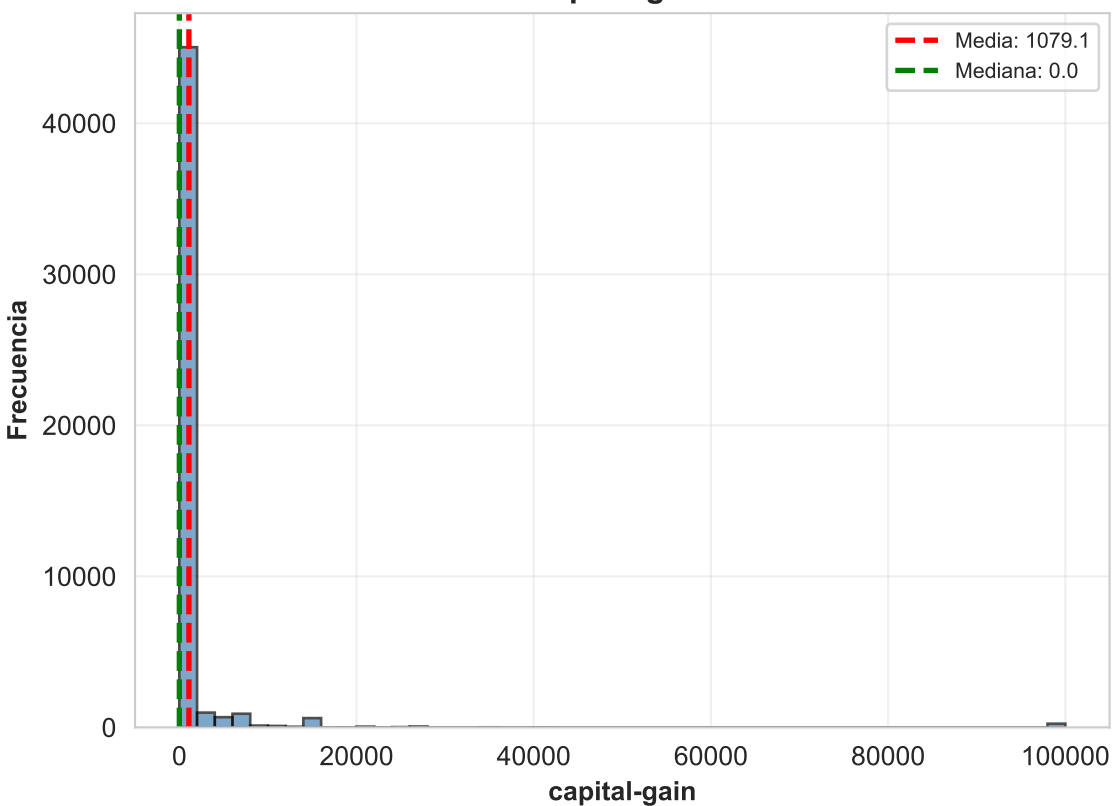
fnlwgt



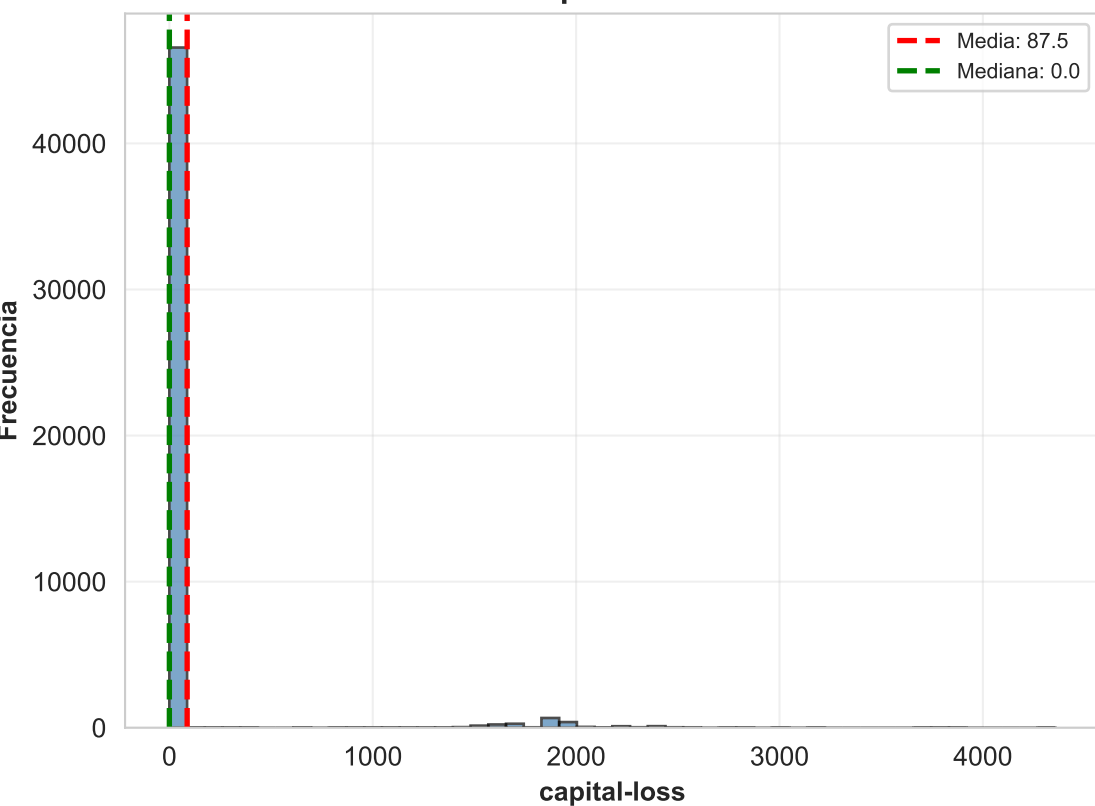
educational-num



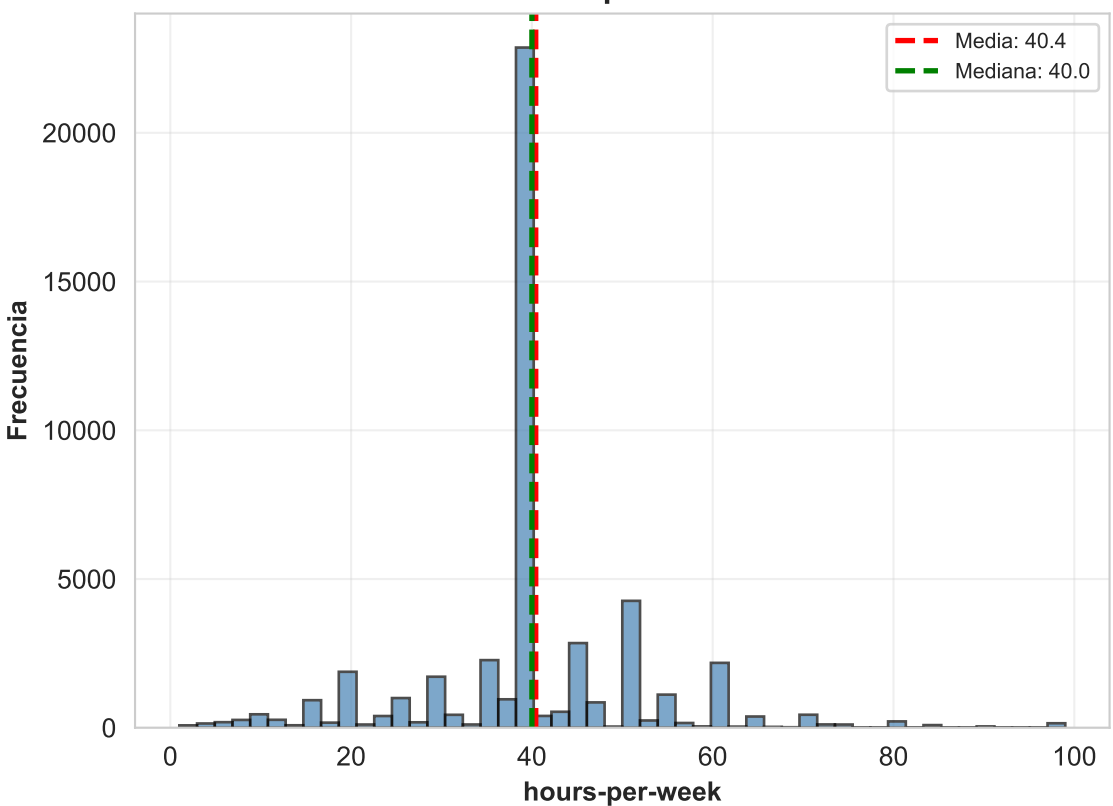
capital-gain



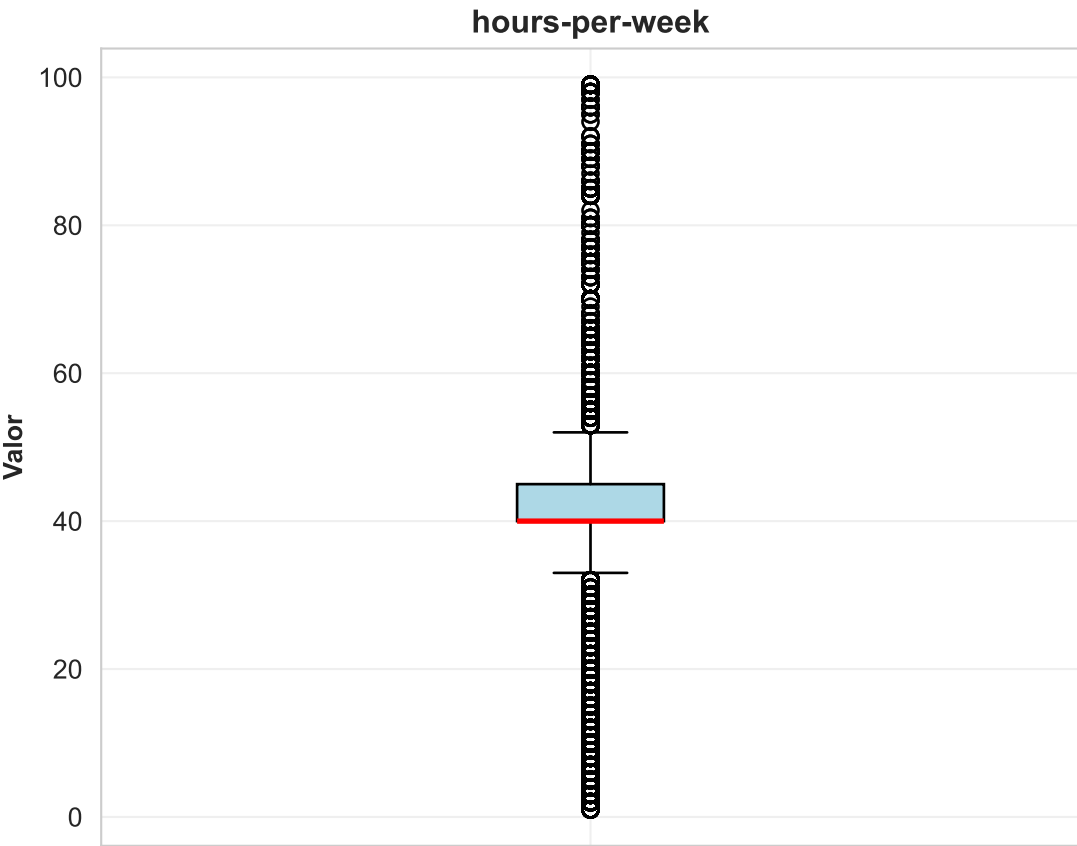
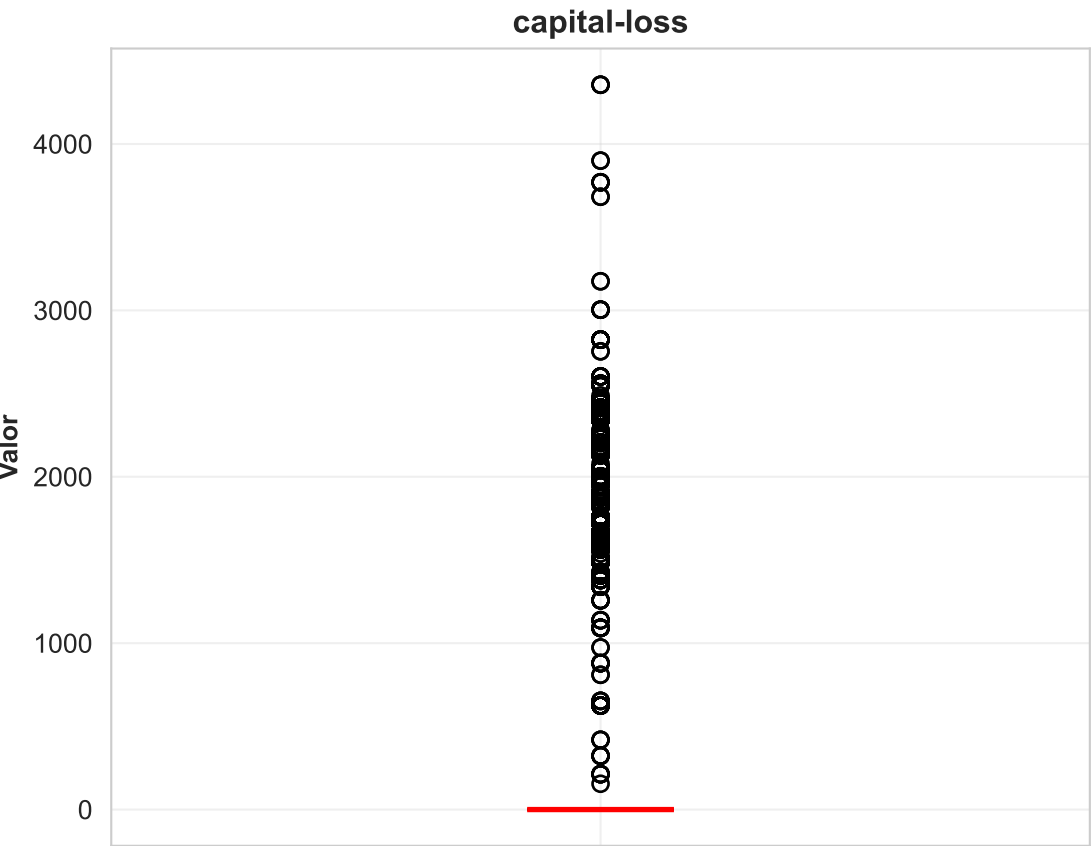
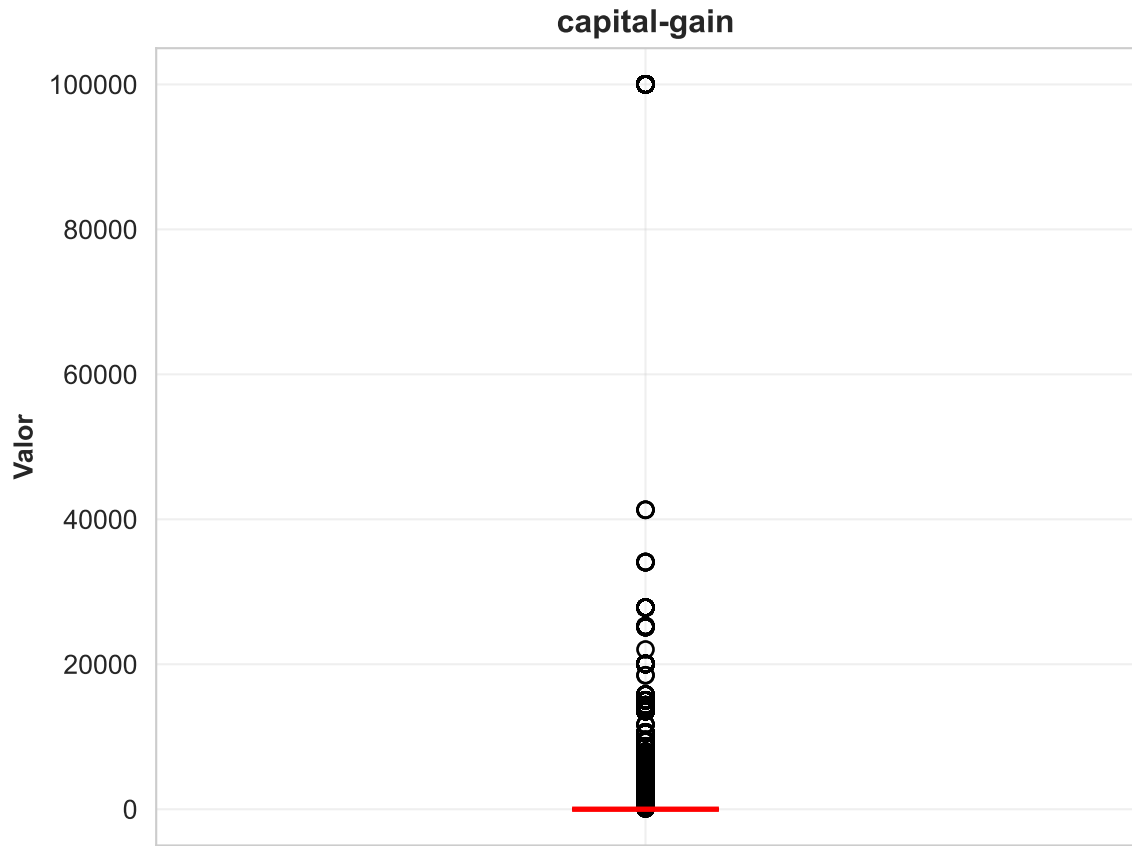
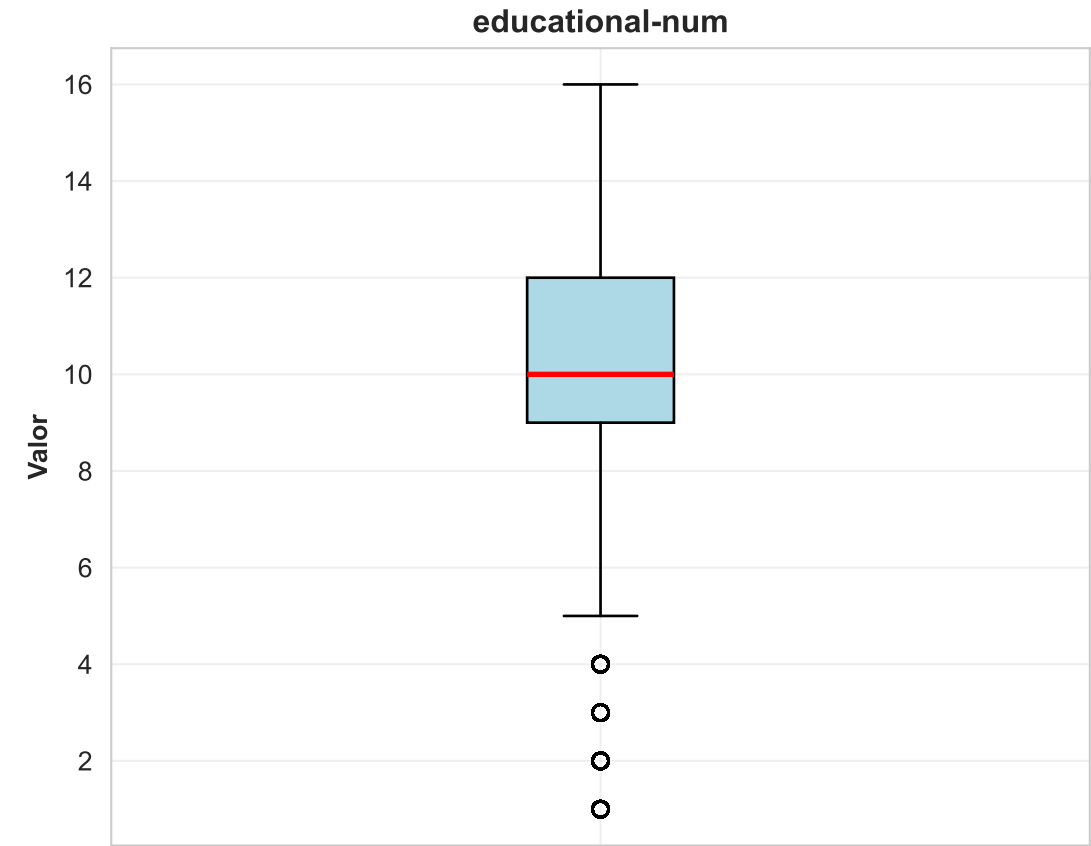
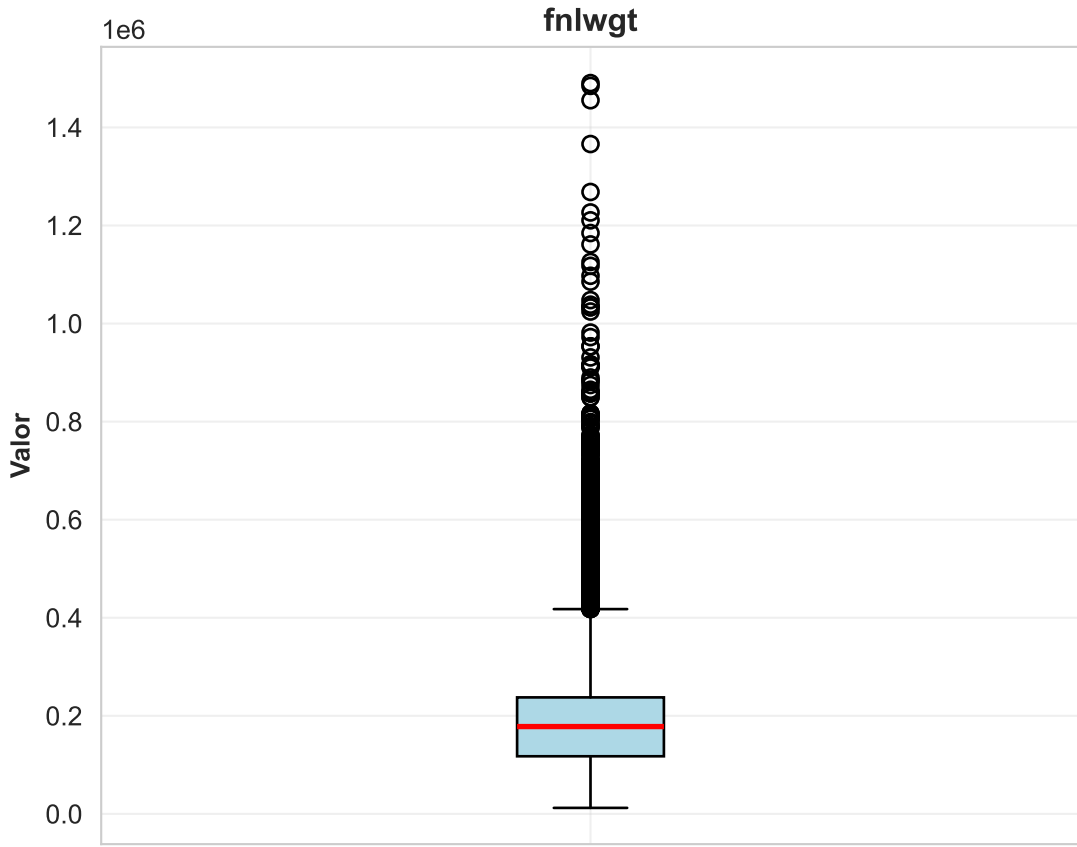
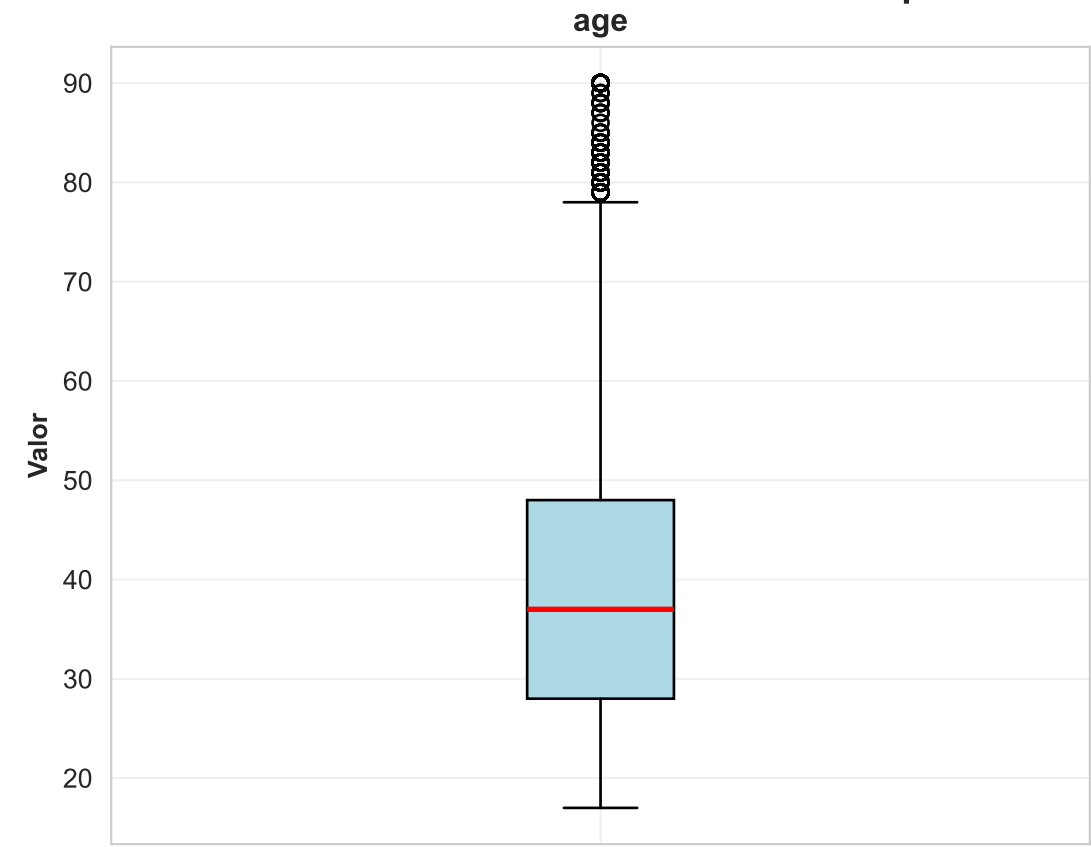
capital-loss



hours-per-week

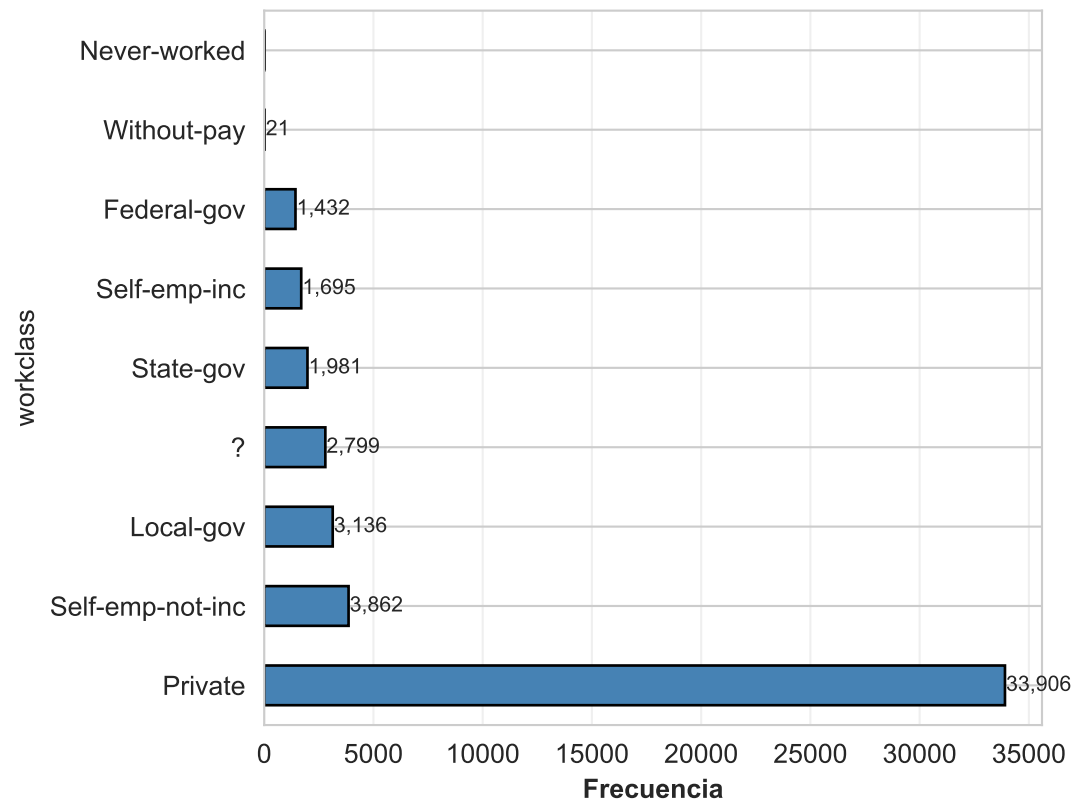


Boxplots - Detección de Outliers

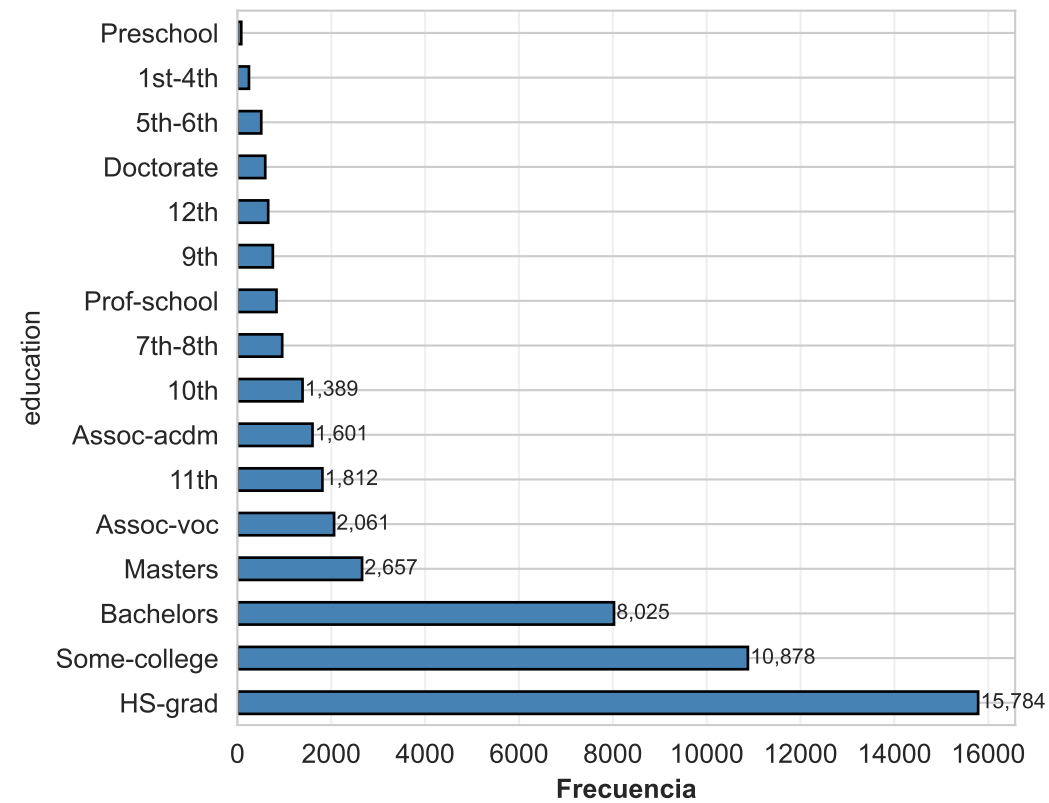


Distribuciones de Variables Categóricas

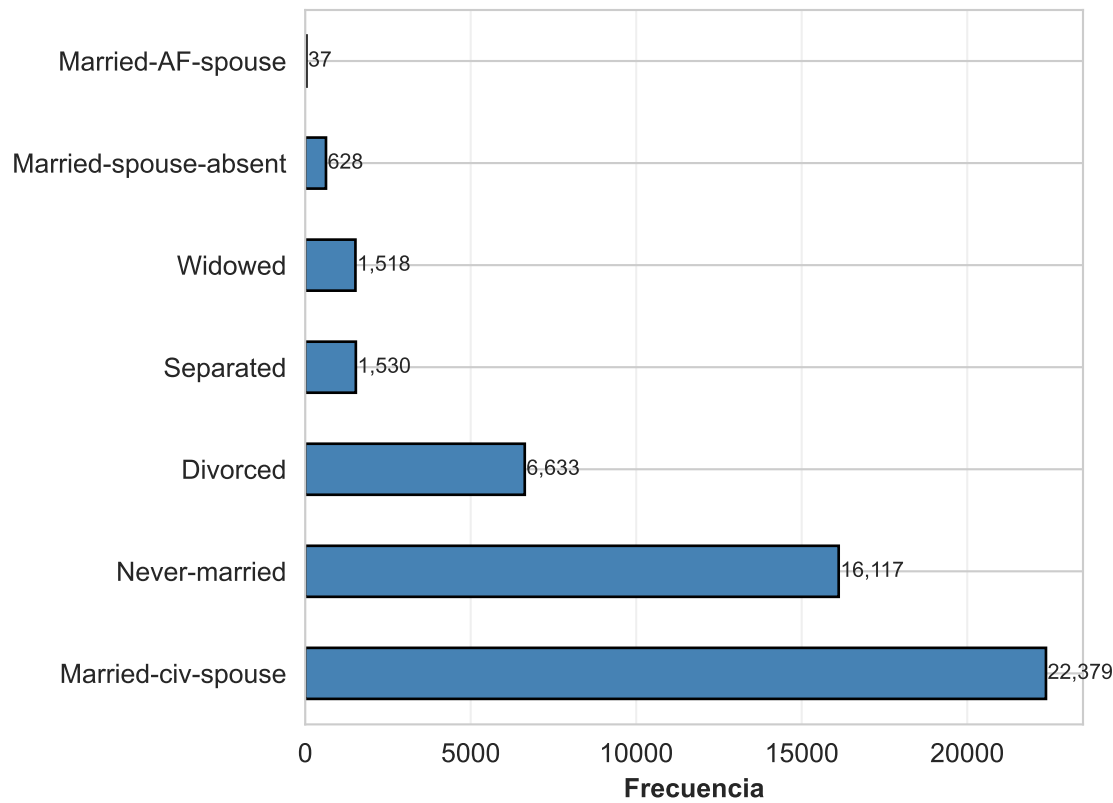
WORKCLASS



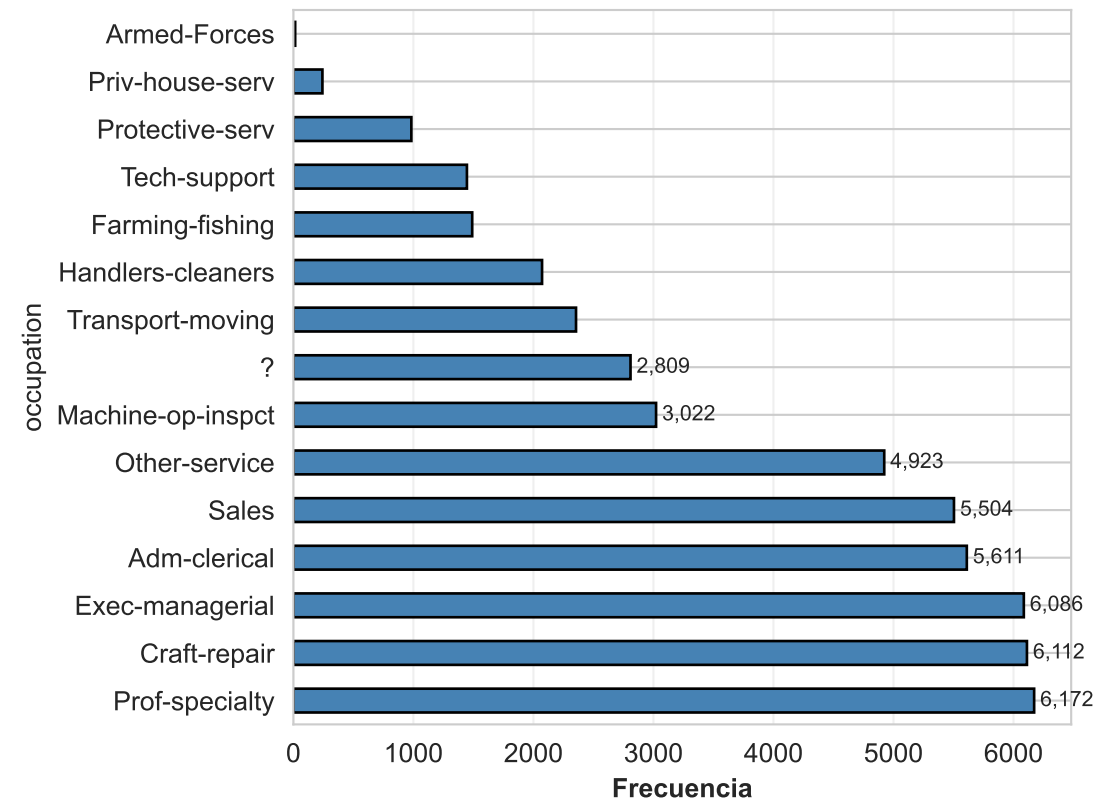
EDUCATION



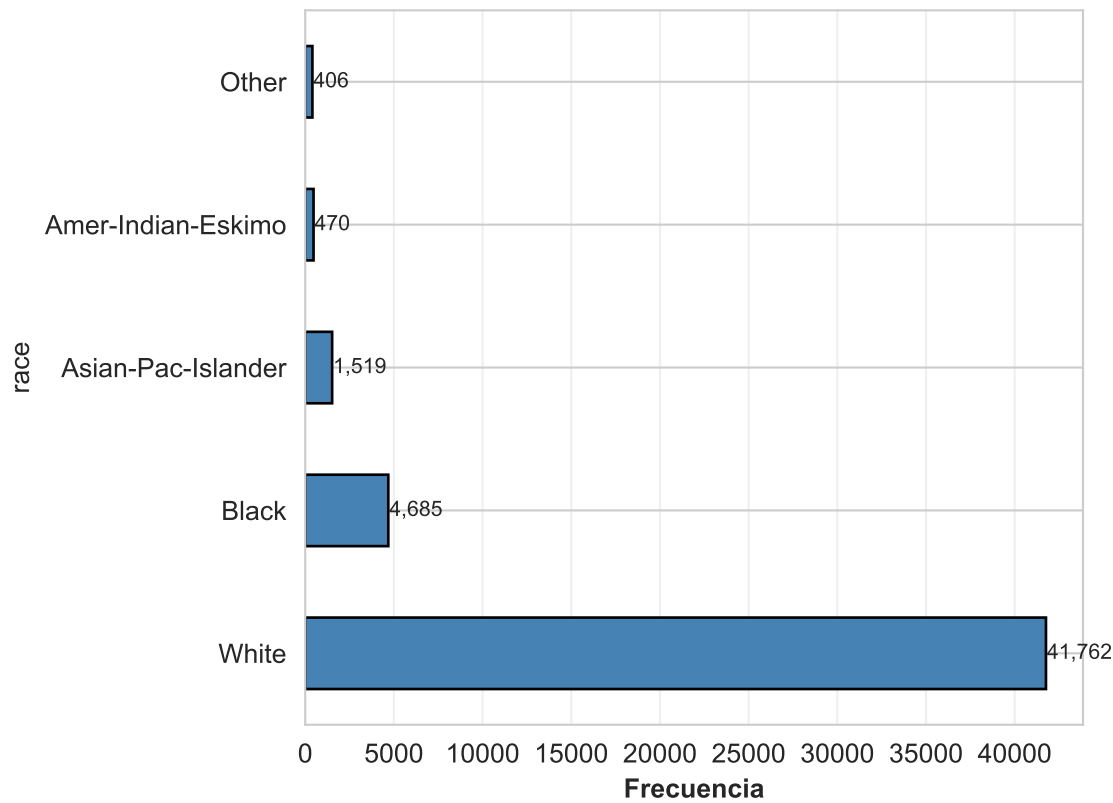
MARITAL-STATUS



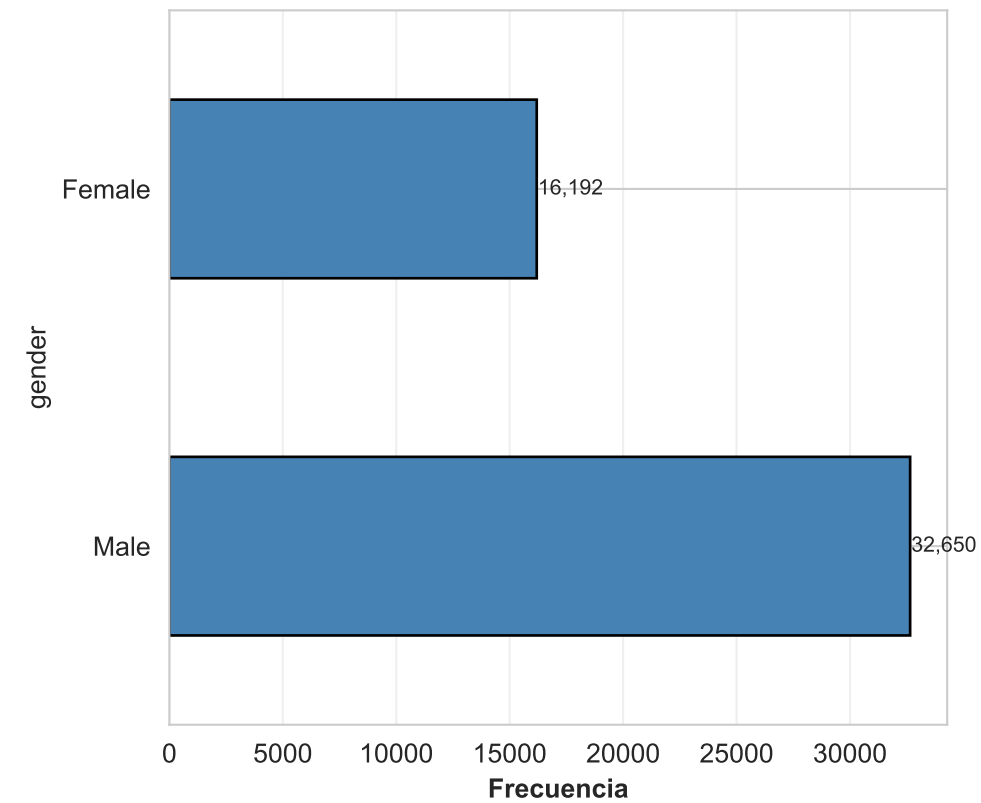
OCCUPATION



RACE

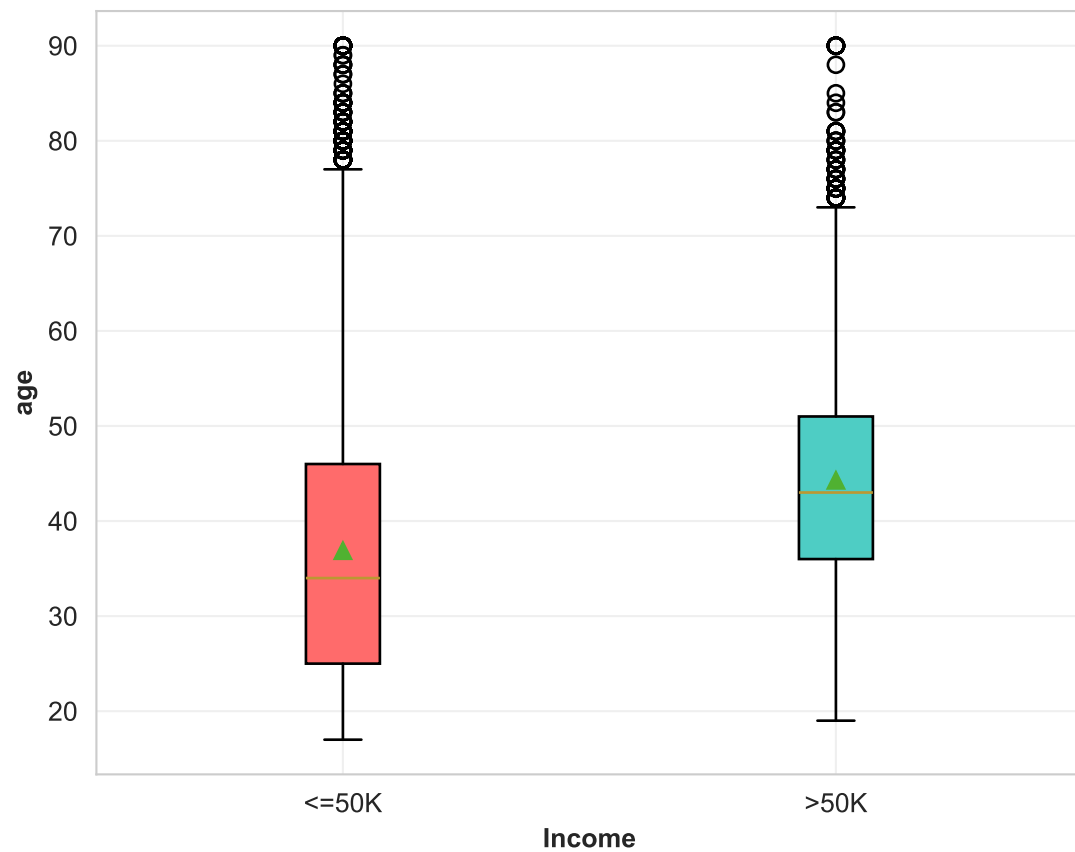


GENDER

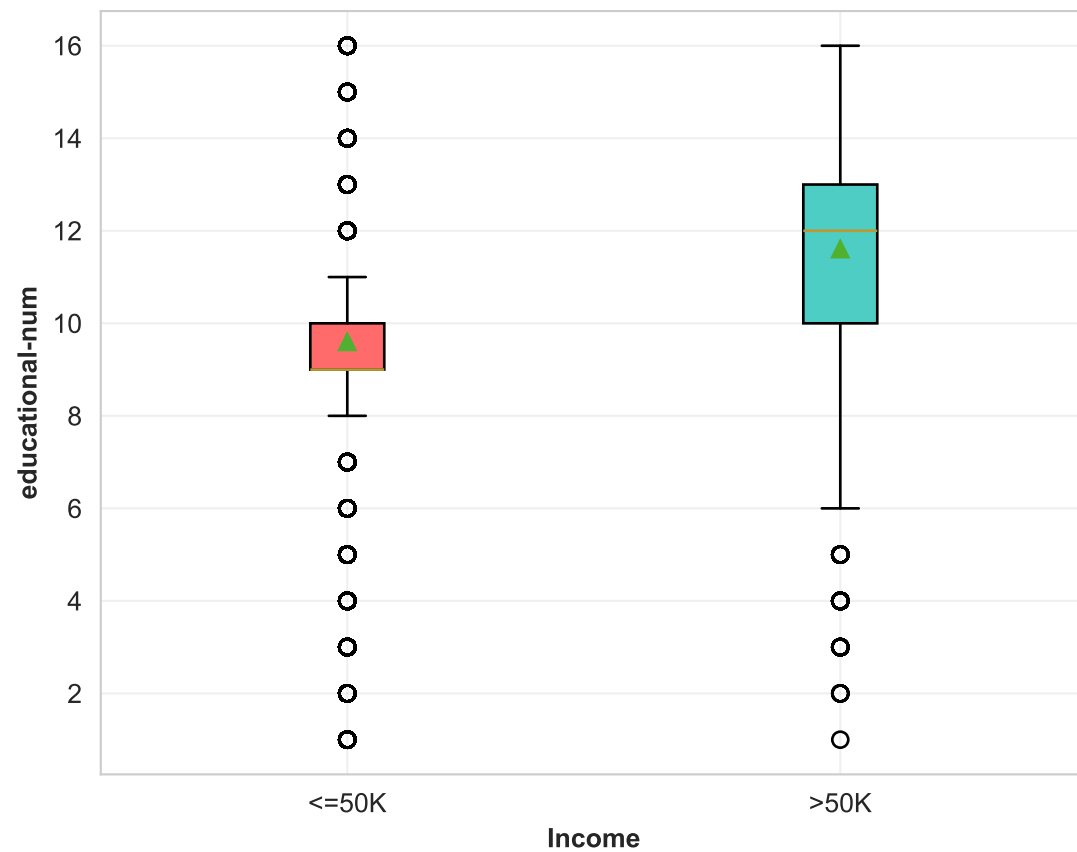


Variables Numéricas por Nivel de Ingreso

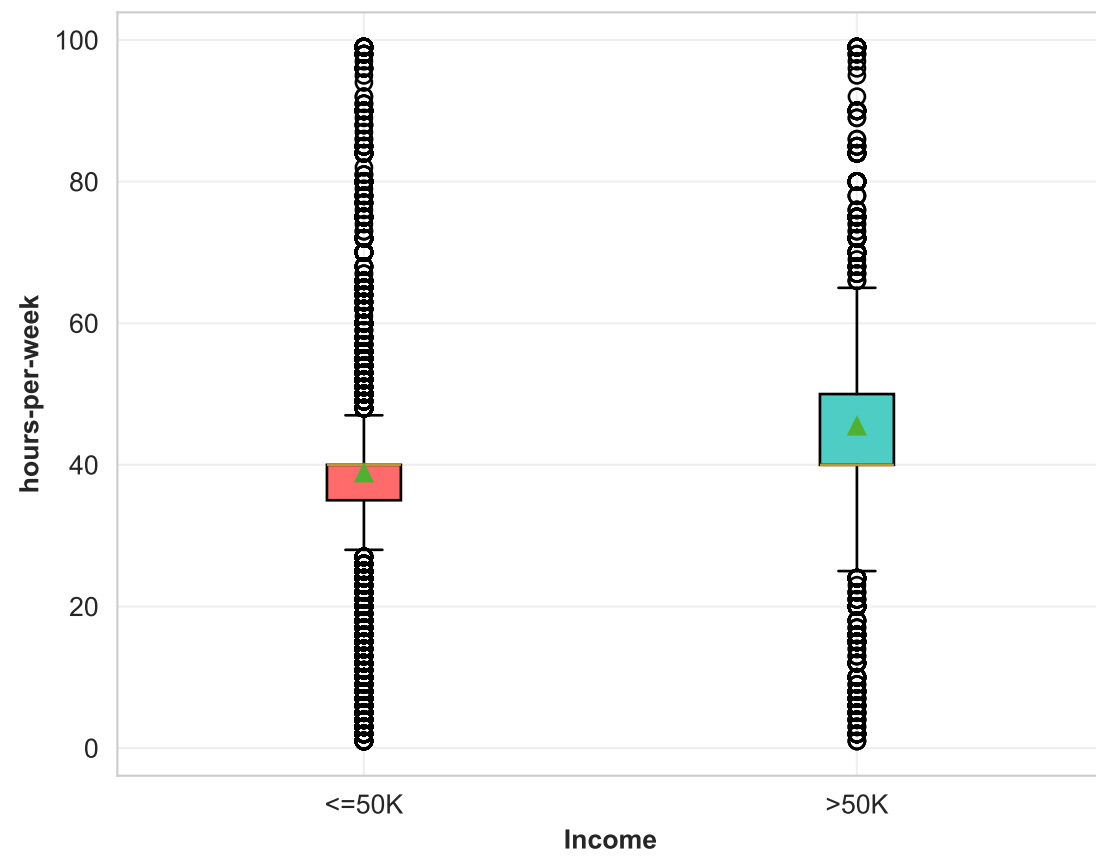
AGE



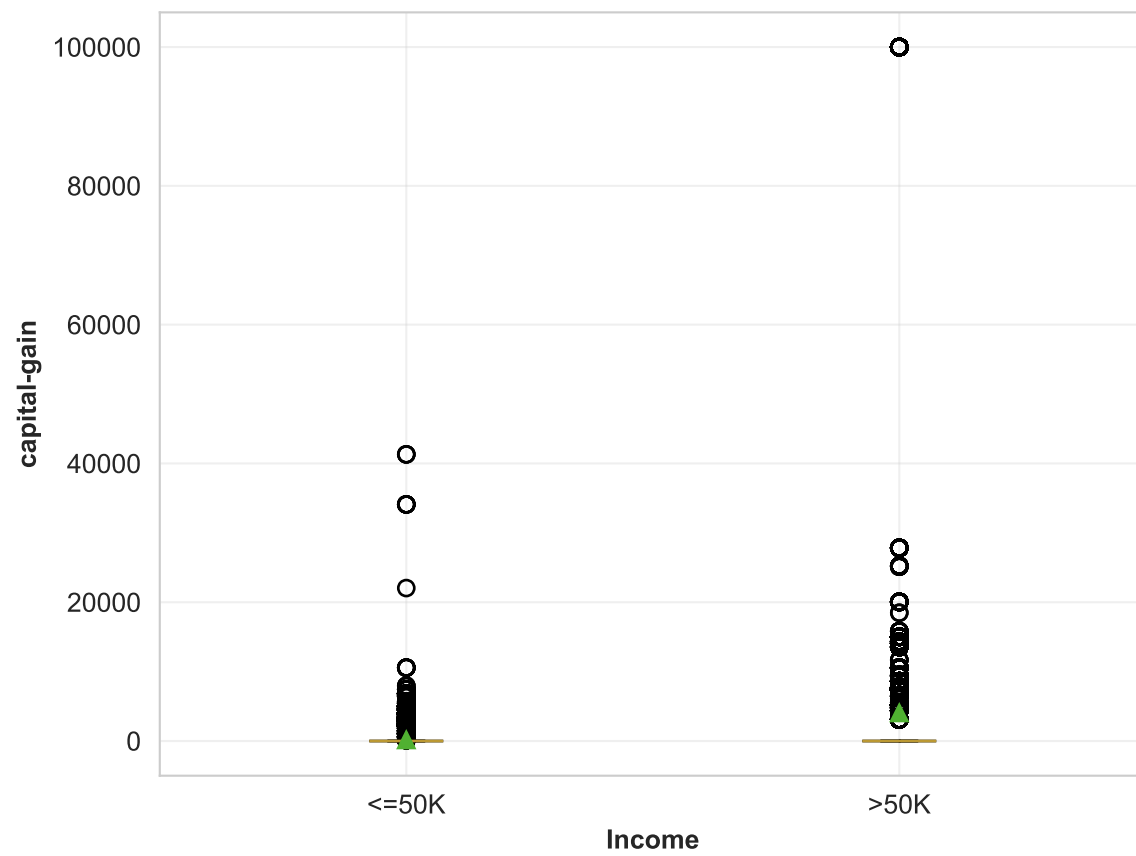
EDUCATIONAL-NUM



HOURS-PER-WEEK

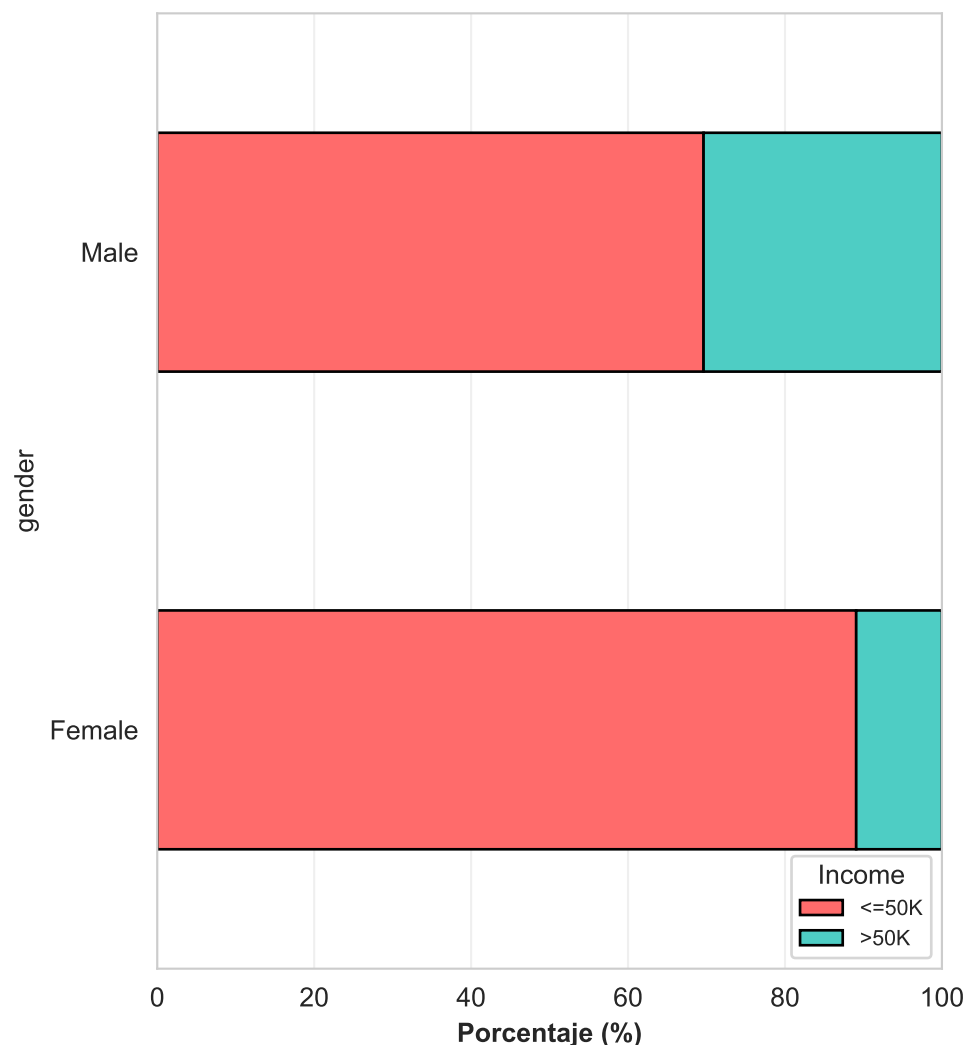


CAPITAL-GAIN

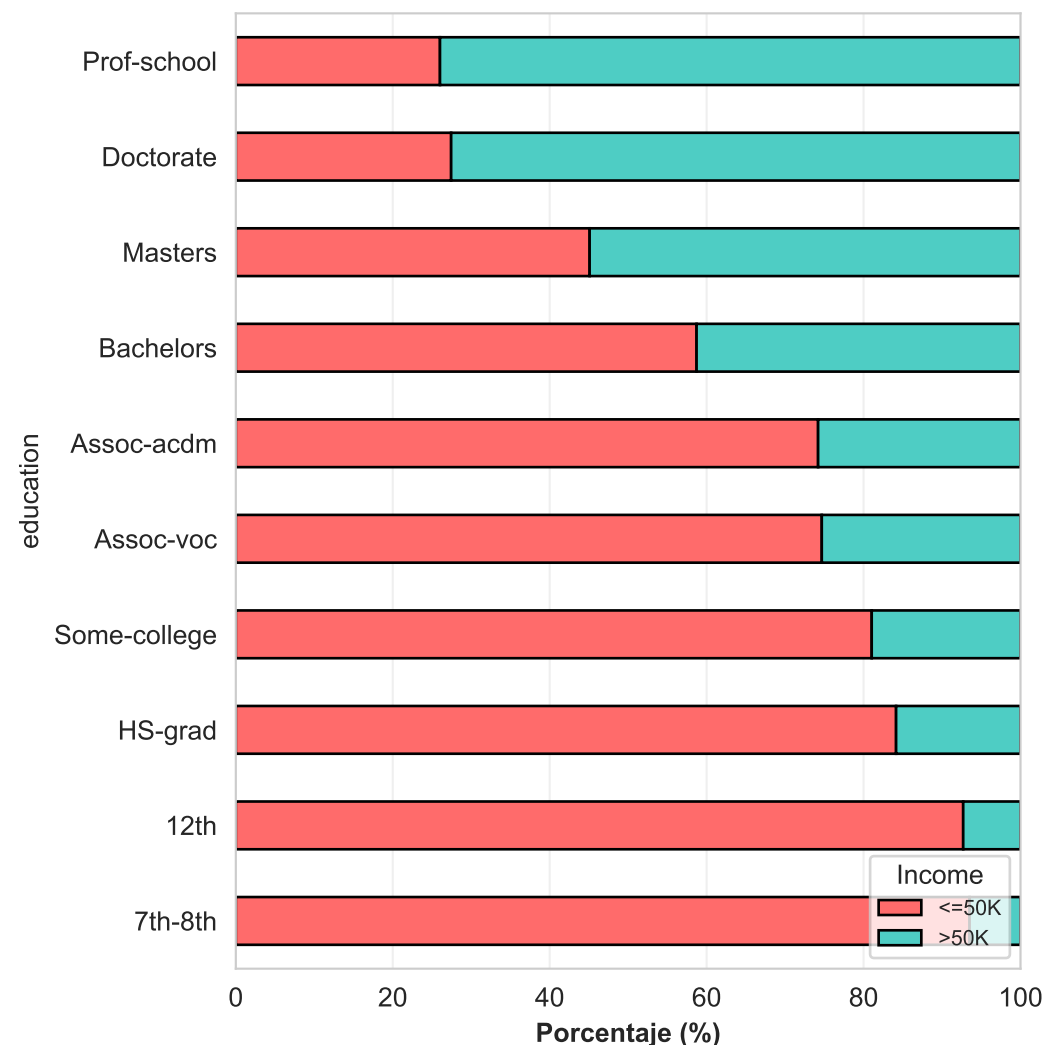


Variables Categóricas vs Income (% Normalizado)

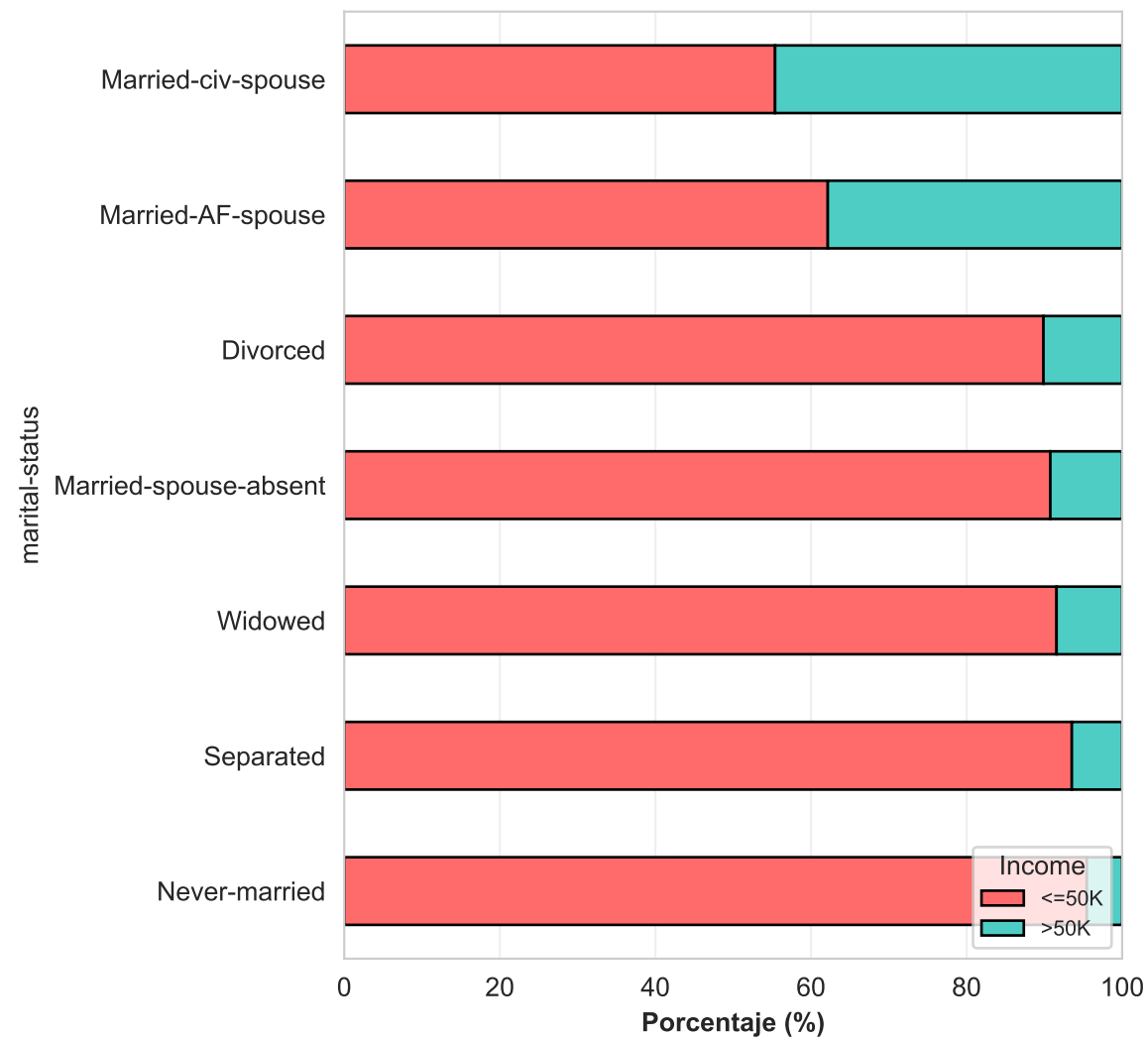
GENDER vs INCOME



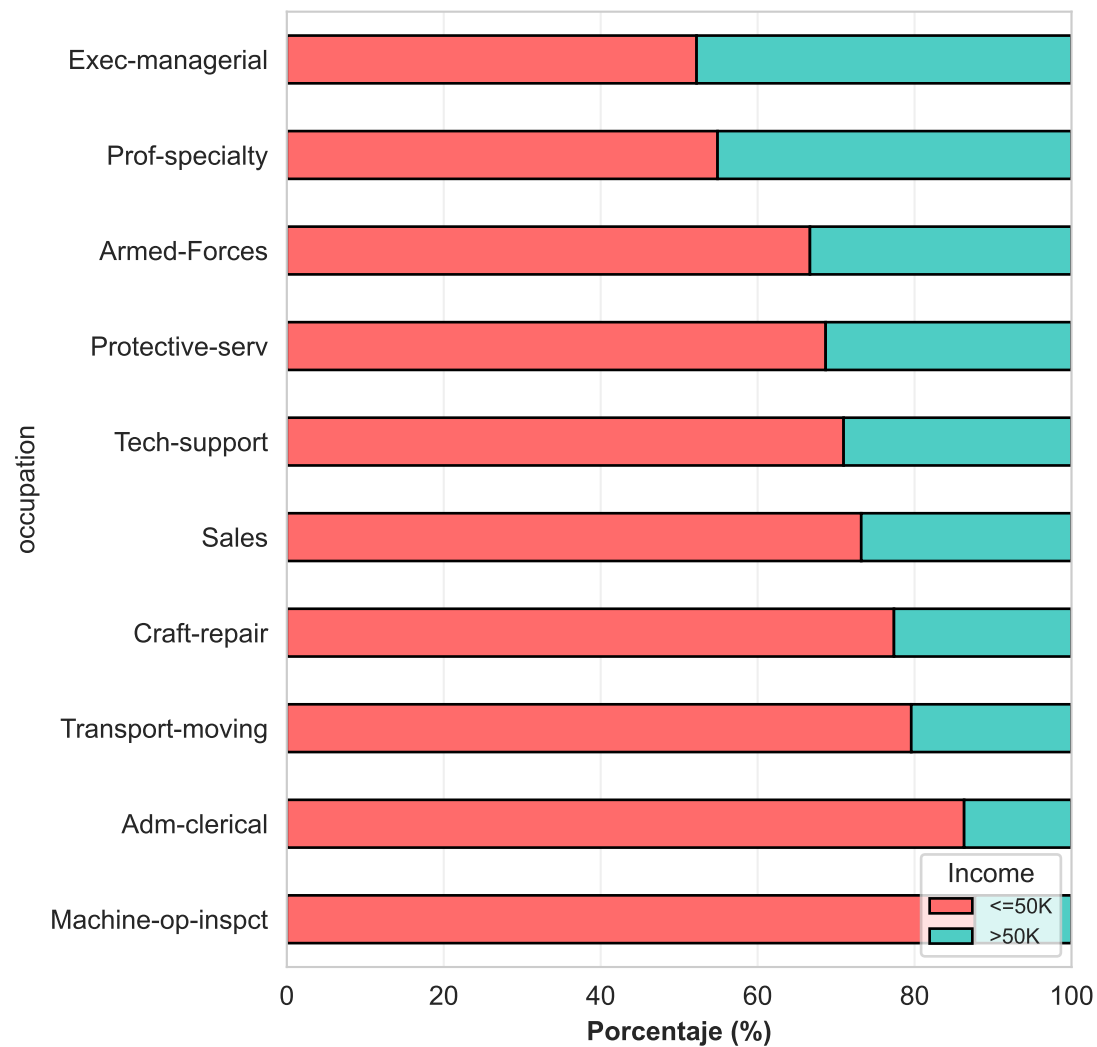
EDUCATION vs INCOME



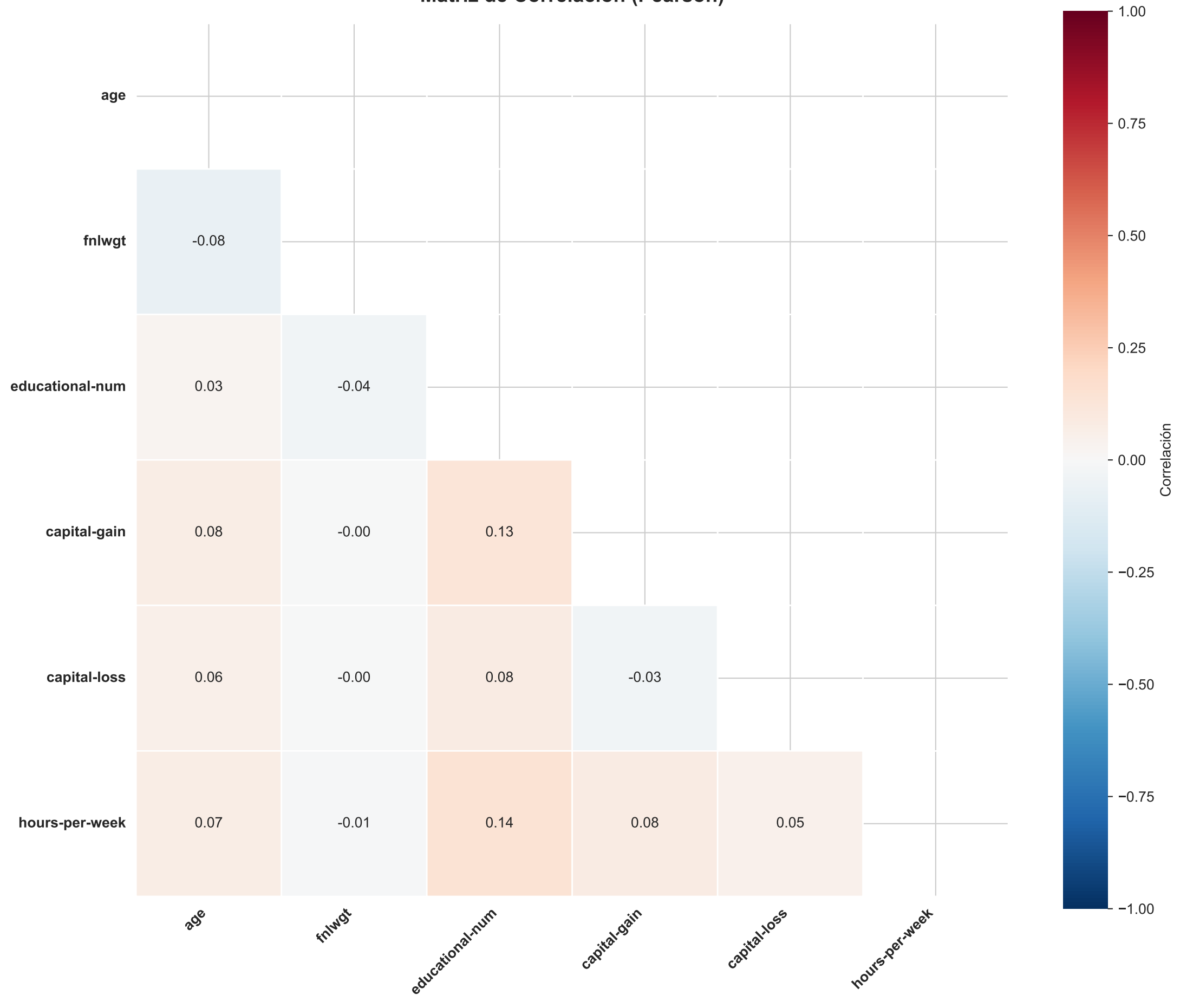
MARITAL-STATUS vs INCOME



OCCUPATION vs INCOME



Matriz de Correlación (Pearson)



HALLAZGOS PRINCIPALES

1. DISTRIBUCIÓN DE INGRESOS
 - Dataset desbalanceado con ratio 3.18:1
 - 37,155 personas ganan ≤50K
 - 11,687 personas ganan >50K
 2. DISPARIDAD DE GÉNERO
 - Masculino: 30.4% ganan >50K
 - Femenino: 10.9% ganan >50K
 - Brecha: 19.5 puntos porcentuales
 3. EDUCACIÓN
 - Prof-school: 74.0%
 - Doctorate: 72.6%
 - Masters: 54.9%
 4. HORAS TRABAJADAS
 - Income >50K: 45.5 horas/semana
 - Income ≤50K: 38.8 horas/semana
 5. DEMOGRAFÍA
 - Edad promedio: 38.6 años
 - Estado civil top: Married-civ-spouse (44.6%)
-

CONCLUSIONES

- La EDUCACIÓN es el factor más determinante para ingresos altos
 - Existe DISPARIDAD DE GÉNERO significativa
 - HORAS TRABAJADAS correlacionan con mayores ingresos
 - ESTADO CIVIL influye en probabilidad de ingresos >50K
 - Variables clave: educational-num, age, hours-per-week
 - Presencia de outliers en capital-gain y capital-loss
-