

Análisis de los datos modelados

Descripción de las variables en la base de datos final

Posterior al proceso de limpieza, estandarización y modelado, se consolidó una base de datos analítica compuesta por 24.963 registros y 26 variables. Estas variables corresponden tanto a información original depurada como a campos derivados producto de transformaciones monetarias, geográficas y de normalización estructural.

A continuación, se presenta la descripción de las variables finales:

Variables laborales y de perfil profesional

Variable: rango_edad

Tipo: Texto categórico

Descripción: Rango de edad reportado por la persona encuestada.

Variable: industria

Tipo: Texto

Descripción: Industria o sector económico donde trabaja la persona encuestada.

Variable: nombre_trabajo

Tipo: Texto

Descripción: Cargo o rol desempeñado dentro de la organización.

Variables de experiencia

Variable: experiencia_general

Tipo: Texto categórico

Descripción: Rango de años de experiencia profesional total acumulada.

Variable: experiencia_campo

Tipo: Texto categórico

Descripción: Rango de años de experiencia específica dentro del campo profesional.

Variables demográficas

Variable: nivel_educacion

Tipo: Texto categórico

Descripción: Nivel educativo más alto alcanzado por la persona encuestada.

Variable: genero

Tipo: Texto categórico

Descripción: Género reportado por la persona encuestada.

Variable: etnia

Tipo: Texto

Descripción: Autoidentificación étnica reportada.

Variables geográficas

Variable: pais

Tipo: Texto

Descripción: País de trabajo estandarizado mediante procesos de homogeneización.

Variable: estado_pais

Tipo: Texto

Descripción: Estado o subdivisión administrativa del país.

Variable: ciudad

Tipo: Texto

Descripción: Ciudad reportada originalmente en el formulario.

Variable: ciudad_clean

Tipo: Texto

Descripción: Ciudad normalizada para análisis geográfico.

Variable: ciudad_imputada_capital

Tipo: Booleano

Descripción: Indica si la ciudad fue imputada utilizando la capital del país.

Variables monetarias originales depuradas

Variable: salario_anual_monto

Tipo: Numérico entero

Descripción: Salario anual limpio en la moneda original reportada.

Variable: compensacion_monto

Tipo: Numérico decimal

Descripción: Compensaciones adicionales limpias en la moneda original.

Variables de moneda

Variable: moneda

Tipo: Texto

Descripción: Moneda reportada inicialmente.

Variable: otra_moneda

Tipo: Texto

Descripción: Moneda ingresada manualmente cuando no estaba en la lista.

Variable: moneda_total

Tipo: Texto

Descripción: Consolidación de moneda y otra_moneda.

Variable: moneda_total_clean

Tipo: Texto

Descripción: Moneda estandarizada para conversión financiera.

Variables de conversión monetaria

Variable: fx_to_usd

Tipo: Numérico decimal

Descripción: Tasa de conversión desde moneda original hacia USD.

Variable: salario_usd

Tipo: Numérico decimal

Descripción: Salario anual convertido a dólares estadounidenses.

Variable: compensacion_usd

Tipo: Numérico decimal

Descripción: Compensaciones adicionales convertidas a dólares.

Variables en pesos colombianos

Variable: salario_cop

Tipo: Numérico decimal

Descripción: Salario anual convertido a pesos colombianos (COP).

Variable: compensacion_cop

Tipo: Numérico decimal

Descripción: Compensaciones adicionales convertidas a COP.

Variable: total_compensacion_cop

Tipo: Numérico decimal

Descripción: Ingreso total anual (salario + compensaciones) en COP.