**EDA + Preprocesamiento desde clean\_raw\_v2 — Obesity Dataset**

**Limpieza aplicada**

• Duplicados eliminados: 0

• Outliers cap (IQR\*3) aplicado a columnas numéricas (winsorization).

• Imputación final: mediana (numéricos) y moda (categóricas).

**EDA**

• Filas: 2087 | Columnas: 17

• Numéricas: 8 | Categóricas: 9

• Resumen numérico exportado a: obesity\_eda\_summary\_v2.csv

**Preprocesamiento**

• One-hot encoding: 12 columnas creadas

• Normalización z-score aplicada a 8 columnas numéricas

• Dataset para modelado: obesity\_model\_ready\_from\_clean.csv

**Reducción de dimensionalidad**

• PCA 2D exportado a: obesity\_pca\_from\_clean.csv

• Varianza explicada: 0.22, 0.19