



Universidad de
los Andes
Colombia

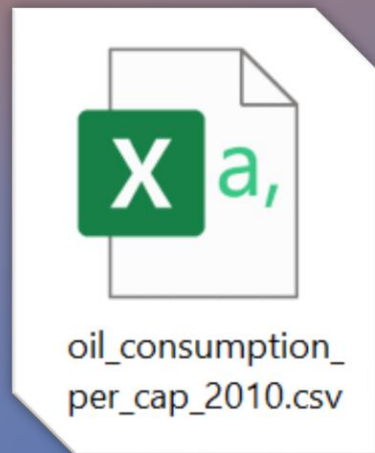
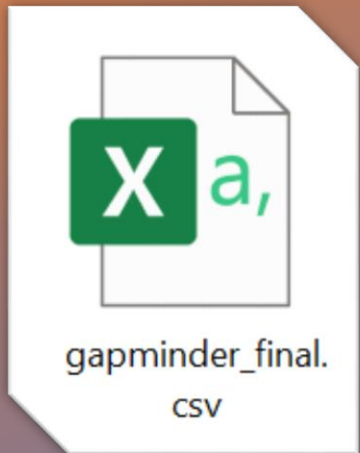
MINE - 4101: Ciencia de Datos Aplicada

Generación de Resultados e Informe

Nicolás Adolfo Castillo Betancourt - na.castillo1@uniandes.edu.co

08 de octubre de 2023

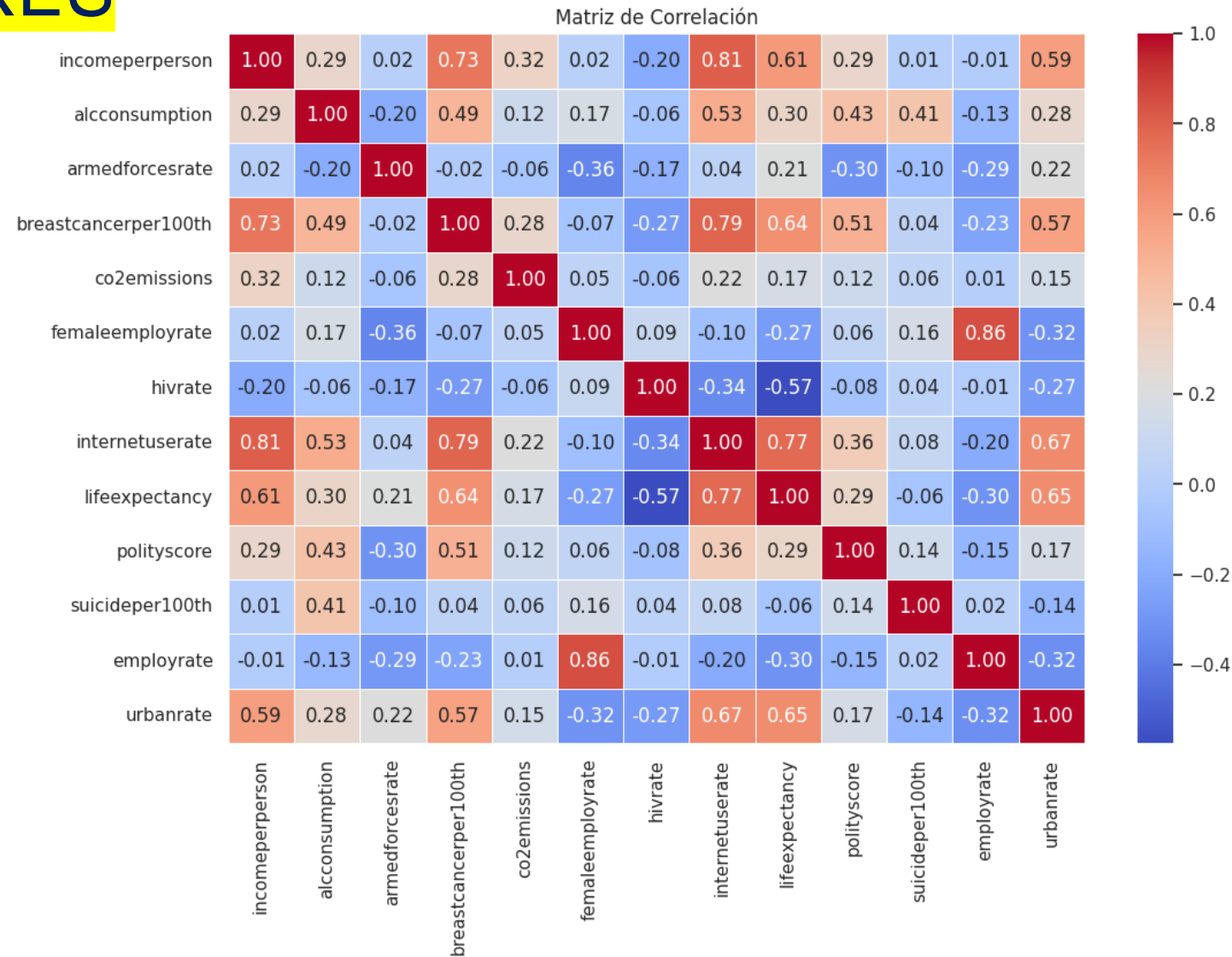
Contexto de Negocio: Apoyo del Banco Mundial a países en desarrollo



gapminder



TOP 5 FEATURES



TOP 5 FEATURES (Relaciones más Importantes)

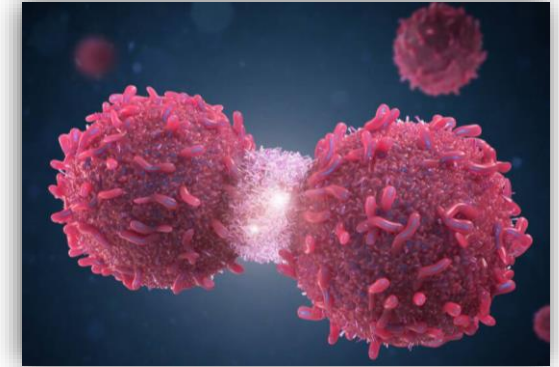
2 - internetuserate



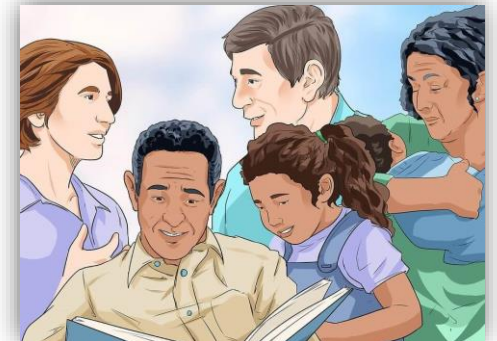
1 - incomeperperson



3 - breastcancerper100th



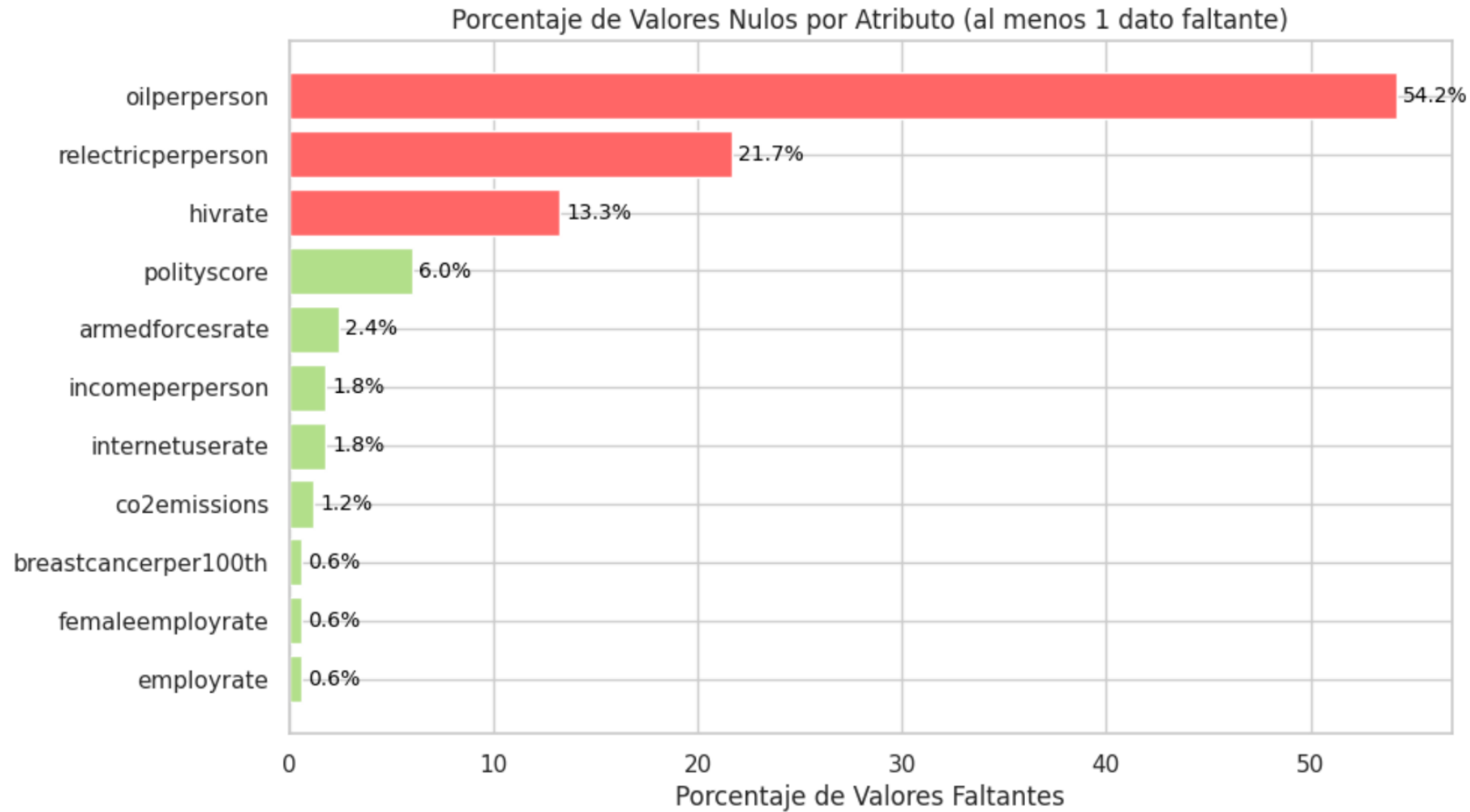
4 - lifeexpectancy



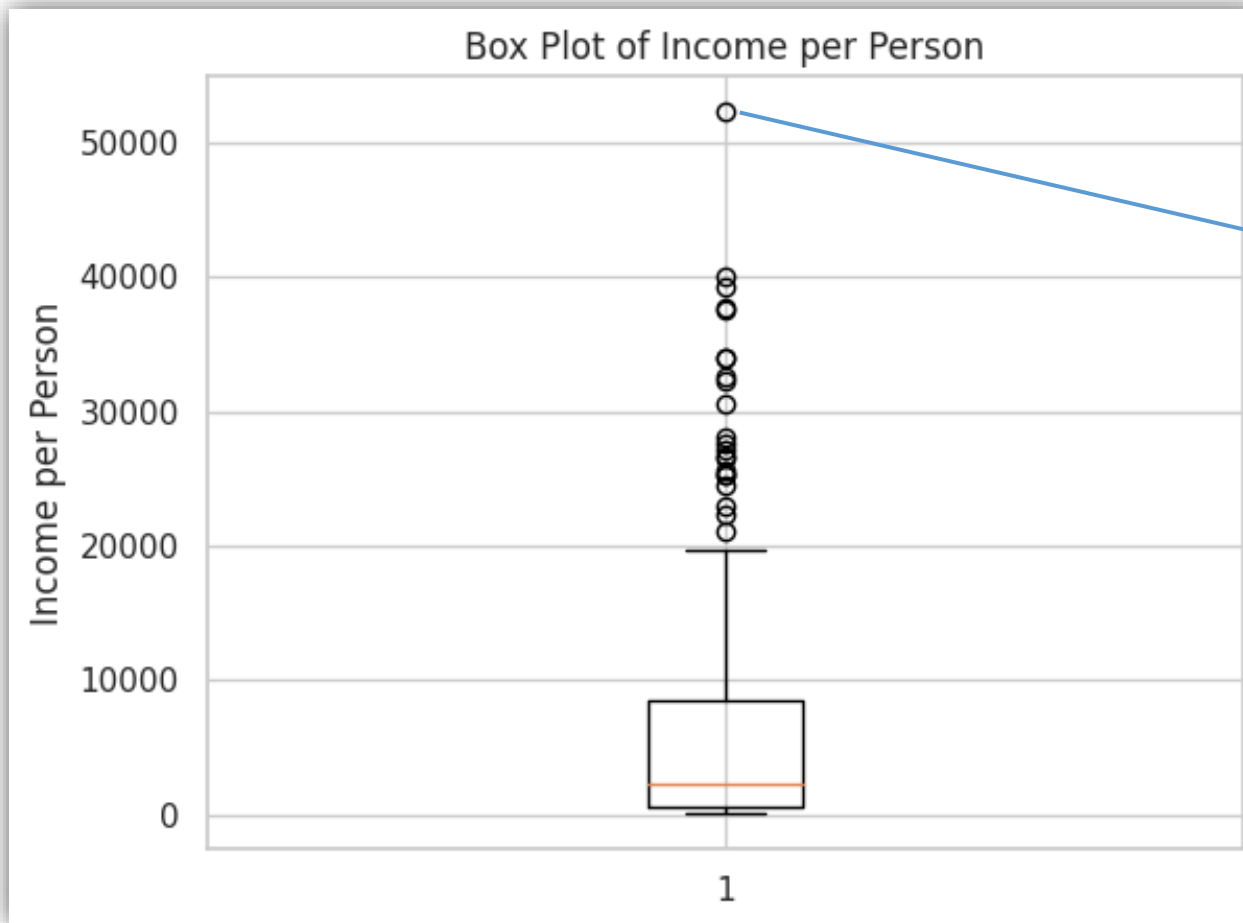
5 - urbanrate



El Caso de Oilperson: Completitud y Descarte



Valor atípico extremo (Luxembourg)



Métricas

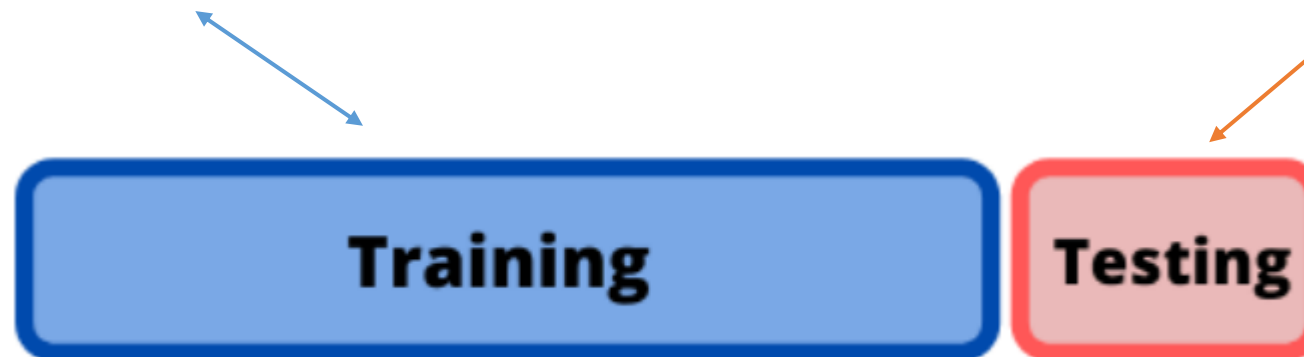
- `model.intercept` -> Valor del término de intercepción: **6784.79**

- **Métricas para el conjunto de Train:**

Error cuadrático medio (MSE): 31295601.99
Error absoluto medio (MAE): 4010.03
Raíz del error cuadrático medio (RMSE): 5594.25
Coeficiente de determinación (R^2): **0.68**

- **Métricas para el conjunto de Test:**

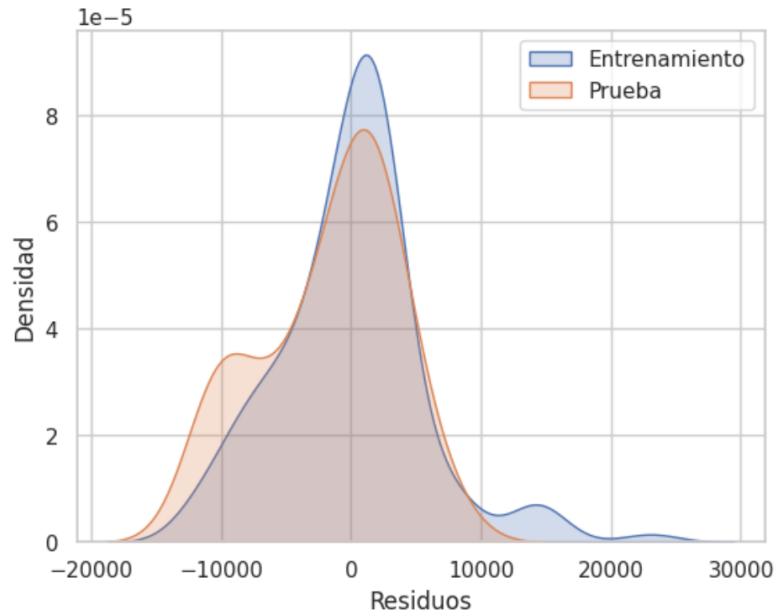
Error cuadrático medio (MSE): 28986478.42
Error absoluto medio (MAE): 4142.36
Raíz del error cuadrático medio (RMSE): 5383.91
Coeficiente de determinación (R^2): **0.65**



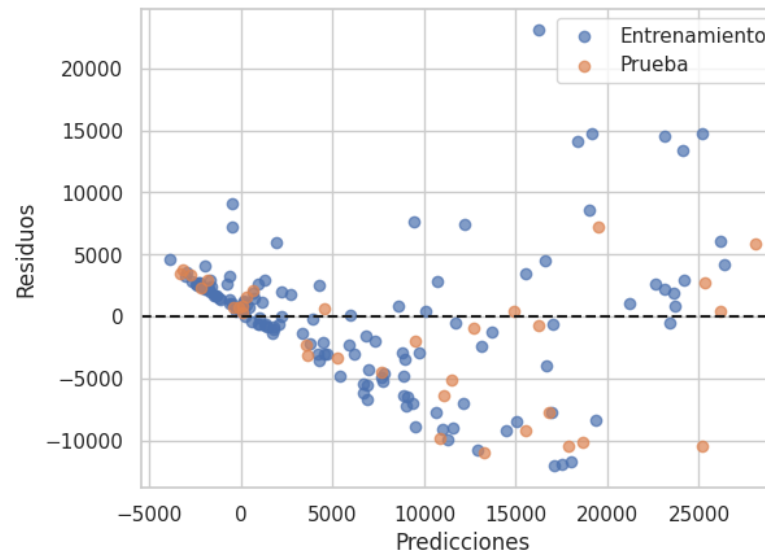
Análisis de resultados

De acuerdo con el valor del coeficiente de determinación (R^2), se puede concluir que el modelo ha logrado capturar aproximadamente el **68%** de la varianza presente en los datos. Este resultado sugiere que el modelo muestra un ajuste razonable a los datos.

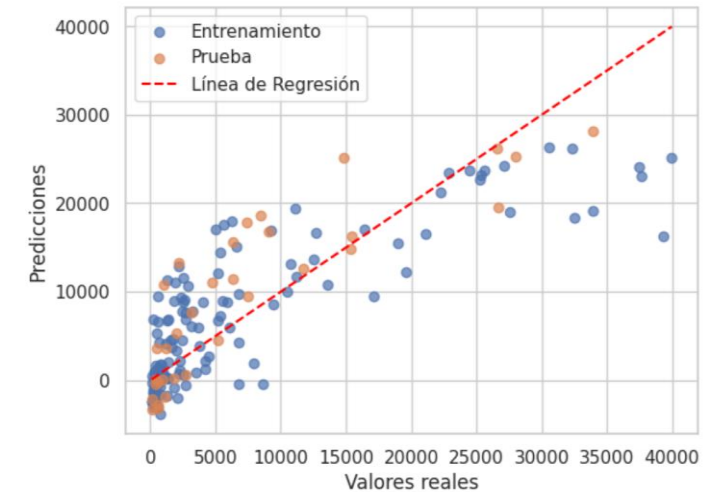
Densidad de Residuos



Scatter Plot



Dispersión con la línea de regresión



Gracias



NICOLAS ADOLFO CASTILLO BETANCOURT