[11pt]article
[breakable]tcolorbox parskip
graphicx caption nocaption format=nocaption,aboveskip=0pt,belowskip=0pt
float figureH xcolor enumerate geometry amsmath amssymb textcomp upquote eurosym
iftex [T1]fontenc alphabeta fontspec unicode-math
fancyvrb grffile [Export]adjustbox max size=0.90.9
hyperref titling longtable booktabs array calc [inline]enumitem [normalem]ulem soul mathrsfs
urlcolorrgb0,.145,.698 linkcolorrgb.71,0.21,0.01 citecolorrgb.12,.54,.11
ansi-blackHTML3E424D ansi-black-intenseHTML282C36 ansi-redHTMLE75C58 ansi-red-intenseHTMLB22B31 ansi-greenHTML00A250 ansi-green-intenseHTML007427 ansi-yellowHTMLDDB62B ansi-yellow-intenseHTMLB27D12 ansi-blueHTML208FFB ansi-blue-intenseHTML0065CA ansi-magentaHTMLD160C4 ansi-magenta-intenseHTMLA03196 ansi-cyanHTML60C6C8 ansi-cyan-intenseHTML258F8F ansi-whiteHTMLC5C1B4 ansi-white-intenseHTMLA1A6B2 ansi-default-inverse-fgHTMLFFFFFF ansi-default-inverse-bgHTML000000
outerrorbackgroundHTMLFFDFDF
HighlightingVerbatimcommandchars=
{}
incolorHTML303F9F outcolorHTMLD84315 cellborderHTMLCFCFCF cellbackgroundHTMLF7F7F7
breaklinks=true, colorlinks=true, urlcolor=urlcolor, linkcolor=linkcolor, citecolor=citecolor,
verbose,tmargin=1in,bmargin=1in,lmargin=1in,rmargin=1in

Análisis cuantitativo evaluación 1

Alejandra Ruiz
Daniel Martinez
Juan Camilo Vergara
Luis Felipe Montenegro

1

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]:
{},codes*=] [rgb]0.24,0.48,0.48# *Imports* [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**pandas** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**pd** [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**seaborn** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**sns** [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**matplotlib**[rgb]0.00,0.00,1.00**.**[rgb]0.00,0.00,1.00**pyplot** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**plt** [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**numpy** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**np** [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**statsmodels**[rgb]0.00,0.00,1.00**.** red[rgb]0.00,0.00,1.00**formula**[rgb]0.00,0.00,1.00**.** red[rgb]0.00,0.00,1.00**api** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**smf** [rgb]0.00,0.50,0.00**import** [rgb]0.00,0.00,1.00**statsmodels**[rgb]0.00,0.00,1.00**.** red[rgb]0.00,0.00,1.00**api** [rgb]0.00,0.50,0.00**as** [rgb]0.00,0.00,1.00**sm** [rgb]0.00,0.50,0.00**from** [rgb]0.00,0.00,1.00**statsmodels**[rgb]0.00,0.00,1.00**.** red[rgb]0.00,0.00,1.00**stats**[rgb]0.00,0.00,1.00**.**[rgb]0.00,0.00,1.00**stattools** [rgb]0.00,0.50,0.00**import** durbin˙watson [rgb]0.00,0.50,0.00**from** [rgb]0.00,0.00,1.00**scipy** [rgb]0.00,0.50,0.00**import** stats

[rgb]0.24,0.48,0.48# *Utils* separator [rgb]0.40,0.40,0.40= [rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13=[rgb]0.73,0.13,0.13’ [rgb]0.40,0.40,0.40* [rgb]0.40,0.40,0.4030 plt[rgb]0.40,0.40,0.40. red˙rcParams[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13text. red˙usetex[rgb]0.73,0.13,0.13’] [rgb]0.40,0.40,0.40= [rgb]0.00,0.50,0.00**True**

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]:
{},codes*=] [rgb]0.24,0.48,0.48# *Dataset* data [rgb]0.40,0.40,0.40= pd[rgb]0.40,0.40,0.40.read˙excel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13. red/datasets/data˙exam1.xlsx[rgb]0.73,0.13,0.13’, sheet˙name[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13data1[rgb]0.73,0.13,0.13’) data[rgb]0.40,0.40,0.40.head()

[breakable, size=fbox, boxrule=.5pt, pad at break*=1mm, opacityfill=0.2]:
{},codes*=] Y X Ind 0 66.199147 12.653765 0 1 44.311301 8.204418 0 2 48.390783 8.768596 0 3 58.087413 16.169568 1 4 60.708671 9.980310 0

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]:
{},codes*=] [rgb]0.24,0.48,0.48# *Categroical data distribution* [rgb]0.00,0.50,0.00print(data[[rgb]0.73,0.13,0.13’][rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’][rgb]0.40,0.40,0.40 red˙value˙counts()) data[[rgb]0.73,0.13,0.13’][rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’][rgb]0.40,0.40,0.40. red˙plot(kind[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13hist[rgb]0.73,0.13,0.13’)

[commandchars=
{},codes*=] Ind 0 800 1 200 Name: count, dtype: int64

[breakable, size=fbox, boxrule=.5pt, pad at break*=1mm, opacityfill=0.2]:
{},codes*=] ¡Axes: ylabel=’Frequency’¿

max size=0.90.9punto$_1$*f**iles/punto*$_{132}$*.png*

Para la variable categórica Ind: - 0: 80% - 1: 20%

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]: {},codes*=] [rgb]0.24,0.48,0.48*# Null values* data[rgb]0.40,0.40,0.40.isnull()[rgb]0.40,0.40,0.40.sum()

[breakable, size=fbox, boxrule=.5pt, pad at break*=1mm, opacityfill=0]: {},codes*=] Y 0 X 0 Ind 0 dtype: int64

No hay valores nulos

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]: {},codes*=] [rgb]0.24,0.48,0.48*# Y distribution* sns[rgb]0.40,0.40,0.40.histplot(data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13'], red kde[rgb]0.40,0.40,0.40=[rgb]0.00,0.50,0.00**True**) plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Distribución de Y[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f**iles/punto*$_{170}$*.png*

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]: {},codes*=] [rgb]0.24,0.48,0.48*# X distribution* sns[rgb]0.40,0.40,0.40.histplot(data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13'], red kde[rgb]0.40,0.40,0.40=[rgb]0.00,0.50,0.00**True**) plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Distribución de X[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f**iles/punto*$_{180}$*.png*

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]: {},codes*=] [rgb]0.24,0.48,0.48*# Relationship between X y Y* sns[rgb]0.40,0.40,0.40.scatterplot(data[rgb]0.40,0.40,0.40=data, x[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', y[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13', hue[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Relación entre X y Y por Categoría de Ind[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f**iles/punto*$_{190}$*.png*

- X y Y siguen una tendencia lineal positiva
- La relación de X y Y por categoría de Ind (0 y 1) es similar, sigue siendo positiva

- Aunque hay una superposición, parece haber una distinción en la ubicación de los puntos azules y naranjas (una agrupación). Esto puede sugerir que la variable 'Ind' tiene un efecto en la relación entre X y Y.

[breakable, size=fbox, boxrule=1pt, pad at breakabovebelow=1mm,colback=cellbackground, colframe=cellborder]: {},codes*=] [rgb]0.24,0.48,0.48# *Pearsons correlation* [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Pearson[rgb]0.73,0.13,0.13') [rgb]0.00,0.50,0.00print(data[[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', red [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13', [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40. redcorr(), [rgb]0.73,0.13,0.13'[rgb]0.67,0.36,0.12**"n**[rgb]0.73,0.13,0.13', separator)

[rgb]0.24,0.48,0.48# *Spearman correlation* [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13 [rgb]0.00,0.50,0.00print(data[[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', red [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13', [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40. redcorr(method[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13spearman[rgb]0.73,0.13,0.13'), red [rgb]0.73,0.13,0.13'[rgb]0.67,0.36,0.12**"n**[rgb]0.73,0.13,0.13', separator)

[rgb]0.24,0.48,0.48# *Kendalls correlation* [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13K [rgb]0.00,0.50,0.00print(data[[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', red [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13', [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40. redcorr(method[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13kendall[rgb]0.73,0.13,0.13'), red [rgb]0.73,0.13,0.13'[rgb]0.67,0.36,0.12**"n**[rgb]0.73,0.13,0.13', separator)

[commandchars= {},codes*=] Pearson X Y Ind X 1.000000 0.832057 0.024069 Y 0. red832057 1.000000 -0.237224 Ind 0.024069 -0.237224 1. red000000 ==========
Spearmans X Y Ind X 1.000000 0.803893 0.028111 Y 0. red803893 1.000000 -0.275985 Ind 0.028111 -0.275985 1.000000 ========== Kendall X Y Ind X 1.000000 0.610515 0.022964 Y 0.610515 1.000000 -0.225454 Ind 0.022964 -0.225454 1.000000 ==========

- Existe una correlación positiva fuerte entre X y Y

[breakable, size=fbox, boxrule=1pt, pad at breakabovebelow=1mm,colback=cellbackground, colframe=cellborder]: {},codes*=] sns[rgb]0.40,0.40,0.40.boxplot(data[rgb]0.40,0.40,0.40=data, x[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13', red y[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Distribución de Y por Categoría de Ind[rgb]0.73,0.13,0.13') plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f iles/punto*$_{11}$3$_0$.*png*

1.0.1 Distribución de los valores de la variable Y para cada una de las categorías de la variable categórica Ind (0 y 1)

- La mediana de la categoría 0 parece ser más alta que la de la categoría 1, lo que sugiere que los valores de Y tienden a ser más altos cuando Ind es 0.
- La categoría 0 tiene un IQR más amplio que la categoría 1, lo que indica una mayor variabilidad en los valores de Y cuando Ind es 0.
- Ambas categorías presentan valores atípicos, lo que indica la presencia de algunos valores de Y que son inusualmente altos o bajos en comparación con el resto de los datos.
- La categoría 0 tiene un rango de datos más amplio en comparación con la categoría 1, sugiriendo que los valores de Y para Ind = 0 varían más que para Ind = 1.

[breakable, size=fbox, boxrule=1pt, pad at breakable*=1mm,colback=cellbackground, colframe=cellborder]:
{},codes*=] [rgb]0.24,0.48,0.48# *Describe X and Y for Ind = 0* description˙ind˙0 [rgb]0.40,0.40,0.40= data[data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13'] [rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', red [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40. reddescribe() [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13"[rgb]0.73,0.13,0.13Description for Ind = 0:[rgb]0.73,0.13,0.13") [rgb]0.00,0.50,0.00print(description˙ind˙0) [rgb]0.00,0.50,0.00print(separator) [rgb]0.24,0.48,0.48# *Describe X and Y for Ind = 1* description˙ind˙1 [rgb]0.40,0.40,0.40= data[data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13'] [rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13', red [rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40. reddescribe() [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13"[rgb]0.73,0.13,0.13Description for Ind = 1:[rgb]0.73,0.13,0.13") [rgb]0.00,0.50,0.00print(description˙ind˙1) [commandchars=
{},codes*=] Description for Ind = 0: X Y count 800.000000 800.000000 mean 9. red931600 49.567376 std 4.055795 23.594228 min -4.263757 -34.894319 25% 7.281930 32.993351 50% 9.873947 49.562809 75% 12.661513 65.509785 max 25. red628678 135.542574 ==========
Description for Ind = 1: X Y count 200.000000 200.000000 mean 10.157892 36.499253 std 2.232061 8.214723 min 3.821281 8.663725 25% 8.865269 31.619189 50% 10.118529 36.079339 75% 11.654149 41.560675 max 16.169568 58.087413

1.0.2 Justificación:

Según la información presentada en el análisis exploratorio sugiere que un modelo de regresión lineal podría ser adecuado para modelar la relación entre X y Y. Basado en la correlación significativa entre X y Y, y asumiendo que la variable Ind también podría influir en Y, sí consideramos posible generar un modelo de regresión lineal para Y incluyendo a Ind como una variable categórica

sin interacción. Esto es posible dado a la fuerte correlación entre X e Y,lo que sugiere una influencia significativa de X en Y.La inclusión de Ind permite evaluar los cambios en el nivel base de Y entre diferentes categorías. Este enfoque mantiene la simplicidad del modelo y la claridad en la interpretación, explorando posibles diferencias entre categorías en una etapa preliminar.

1.1 Modelo de regresión lineal

[breakable, size=fbox, boxrule=1pt, pad at breakable=1mm,colback=cellbackground, colframe=cellborder]: {},codes*=] model [rgb]0.40,0.40,0.40= smf[rgb]0.40,0.40,0.40. red ols([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y ~ X + C(Ind)[rgb]0.73,0.13,0.13', data[rgb]0.40,0.40,0.40=data)[rgb]0.40,0.40,0.40. red fit() [rgb]0.00,0.50,0.00print(model[rgb]0.40,0.40,0.40.summary()) [commandchars= {},codes*=] OLS Regression Results ==========
Dep. Variable: Y R-squared: 0.759 Model: OLS Adj. R-squared: 0.758 Method: Least Squares F-statistic: 1566. Date: Tue, 02 Apr 2024 Prob (F-statistic): 2.25e-308 Time: 16:19:07 Log-Likelihood: -3801.1 No. Observations: 1000 AIC: 7608. Df Residuals: 997 BIC: 7623. Df Model: 2 Covariance Type: nonrobust
==========
coef std err t P¿—t— [0.025 0.975] ———— ———- Intercept 0.7873 0.984 0.800 0. red424 -1.143 2.718 C(Ind)[T.1] -14.1796 0.858 -16.535 0. red000 -15.862 -12.497 X 4.9116 0.091 53.848 0.000 4.733 5.091
==========
Omnibus: 2.517 Durbin-Watson: 1.997 Prob(Omnibus): 0.284 Jarque-Bera (JB): 2.440 Skew: 0.078 Prob(JB): 0.295 Kurtosis: 3.185 Cond. No. 31.7
==========

Notes: [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

1.2 Análisis de Resultados del Modelo de Regresión Lineal

1.2.1 Estadísticas del Modelo

- **R-cuadrado (R^2)**: 0.759
  - Indica que el 75.9% de la variabilidad de `Y` puede ser explicada por las variables `X` e `Ind`.
- **R-cuadrado ajustado**: 0.758
  - Muestra que el modelo ajusta bien sin ser penalizado significativamente por incluir variables adicionales.
- **F-estadístico**: 1566
  - Sugiere que hay una relación lineal significativa, con un valor-P cercano a 0.

1.2.2 Coeficientes del Modelo

- **Intercepto**: 0.7873 (p-valor: 0.424)
  - No estadísticamente significativo, indicando que el valor esperado de `Y` cuando `X` es 0 y `Ind` es 0, es cercano a 0.7873.
- **C(Ind)[T.1]**: -14.1796 (p-valor: prácticamente 0)
  - Significativo, sugiriendo que `Y` disminuye en promedio 14.1796 unidades cuando `Ind` cambia de 0 a 1, manteniendo `X` constante.
- **X**: 4.9116 (p-valor: prácticamente 0)
  - Muy significativo, indicando que por cada unidad que incrementa `X`, `Y` aumenta en 4.9116 unidades.

1.2.3 Diagnóstico del Modelo

- **Durbin-Watson**: 1.997
  - Implica que no hay evidencia de autocorrelación en los residuos del modelo.
- **Pruebas de Normalidad**:
  - **Omnibus**: 2.517 (p-valor: 0.284)
  - **Jarque-Bera (JB)**: 2.440 (p-valor: 0.295)
  - Ambas pruebas indican que no hay desviaciones significativas de la normalidad en los residuos.

2 Visualización del modelo en el diagrama de disperción

2.0.1 Ind = 0

$$Y_i = \beta_{Intercept} + \beta_X X_i$$

Ind = 1

$$Y_i = \beta_{Intercept} + \beta_{C(Ind)[T.1]} + \beta_X X_i$$

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm,colback=cellbackground, colframe=cellborder] [commandchars={},codes*=] In [31]:
x˙vals [rgb]0.40,0.40,0.40= np[rgb]0.40,0.40,0.40.
redlinspace(data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40.
redmin(), data[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13']][rgb]0.40,0.40,0.40.
redmax(), [rgb]0.40,0.40,0.40100)

[rgb]0.24,0.48,0.48# *Predict Y for Ind=0 and Ind=1 using the fitted model*
[rgb]0.24,0.48,0.48# *model.params['X'] --¿ Line slope*

[rgb]0.24,0.48,0.48# *Creating the prediction lines* [rgb]0.24,0.48,0.48# *Y = b0 + (b1 * X)* pred˙Y˙ind˙0 [rgb]0.40,0.40,0.40= model[rgb]0.40,0.40,0.40.
red params[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Intercept[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40+ model[rgb]0.40,0.40,0.40.params[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0
[rgb]0.40,0.40,0.40* x˙vals [rgb]0.24,0.48,0.48# *Y = b0 + (b1 * X) + (change associated with Ind=1)*
pred˙Y˙ind˙1 [rgb]0.40,0.40,0.40= (model[rgb]0.40,0.40,0.40.
red params[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Intercept[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40+ model[rgb]0.40,0.40,0.40.params[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13C(Ind)[T.
red 1][rgb]0.73,0.13,0.13’] [rgb]0.40,0.40,0.40+ model[rgb]0.40,0.40,0.40.
red params[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40* x˙vals)
[rgb]0.24,0.48,0.48# *Create a figure with 3 subplots* fig, axes [rgb]0.40,0.40,0.40= plt[rgb]0.40,0.40,0.40.subplots([rgb]0.40,0.40,0.401,
red [rgb]0.40,0.40,0.403, figsize[rgb]0.40,0.40,0.40=([rgb]0.40,0.40,0.4015,
[rgb]0.40,0.40,0.405))
title˙no˙interaction˙ind0 [rgb]0.40,0.40,0.40=
[rgb]0.73,0.13,0.13r[rgb]0.73,0.13,0.13”[rgb]0.73,0.13,0.13($Ind = 0$): $Y =
[rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47–**0″**[rgb]0.73,0.13,0.13
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47–**X″**[rgb]0.73,0.13,0.13X$[rgb]0.73,0.13,0.13
title˙no˙interaction˙ind1 [rgb]0.40,0.40,0.40= [rgb]0.73,0.13,0.13r[rgb]0.73,0.13,0.13”[rgb]0.73,0.13,0.13($Ind
= 1$): $Y = ([rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47–**0″**[rgb]0.73,0.13,0.13
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.73,0.13,0.13–[rgb]0.73,0.13,0.13Ind=1″)
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47–**X″**[rgb]0.73,0.13,0.13X$[rgb]0.73,0.13,0.13
[rgb]0.24,0.48,0.48# *Plot 1: Scatter plot of data with Ind = 0* axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
red scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red5, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.plot(x˙vals, pred˙Y˙ind˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13black[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
red set˙title(title˙no˙interaction˙ind0) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
red set˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
[rgb]0.24,0.48,0.48# *Plot 2: Scatter plot of data with Ind = 1* axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
red scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],

red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 1[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red5, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.plot(x˙vals, pred˙Y˙ind˙1,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13black[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙title(title˙no˙interaction˙ind1) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
[rgb]0.24,0.48,0.48# *Plot 3: Regression lines for Ind =
0 and Ind = 1* axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redscatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red5, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 1[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red5, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.plot(x˙vals, pred˙Y˙ind˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión Ind
= 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.plot(x˙vals,
pred˙Y˙ind˙1, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redset˙title([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind = 0 y Ind
= 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
[rgb]0.24,0.48,0.48# *Adjust the spacing between subplots*

plt[rgb]0.40,0.40,0.40.tight˙layout()
[rgb]0.24,0.48,0.48# *Show the plots* plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f iles/punto*$_{12}$1$_0$.*png*

2.0.2 Análisis gráfico

Realizando un análisis a lo presentado por las gráficas anteriores, encontramos que la pendiente parece no describir adecuadamente el comportamiento de los datos. Por lo que realizaremos a continuación una prueba de interacción, es decir, que asumiremos que si hay interacción entre las variables X e Ind. Esto nos permitirá observar si la pendiente de X cambia cuando Ind es 1.

2.1 Prueba de interacción

[breakable, size=fbox, boxrule=1pt, pad at breakable=1mm,colback=cellbackground, colframe=cellborder]: {},codes*=] model˙with˙interaction [rgb]0.40,0.40,0.40= smf[rgb]0.40,0.40,0.40.ols([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Y ˜ X * C(Ind)[rgb]0.73,0.13,0.13', data[rgb]0.40,0.40,0.40=data)[rgb]0.40,0.40,0.40.fit() [rgb]0.00,0.50,0.00print(model˙with˙interaction[rgb]0.40,0.40,0.40.summary())
[commandchars=
{},codes*=] OLS Regression Results ==========
Dep. Variable: Y R-squared: 0.765 Model: OLS Adj. R-squared: 0.764 Method: Least Squares F-statistic: 1081. Date: Tue, 02 Apr 2024 Prob (F-statistic): 1.34e-312 Time: 16:19:08 Log-Likelihood: -3787.5 No. Observations: 1000 AIC: 7583. Df Residuals: 996 BIC: 7603. Df Model: 3 Covariance Type: nonrobust
==========
= coef std err t P¿—t— [0.025 0.975] ———— - Intercept -0.4991 1.001 -0.498 0.618 -2.464 1.466 C(Ind)[T.1] 4.5491 3.674 1.238 0.216 -2.661 11.759 X 5.0411 0.093 53.997 0. red000 4.858 5.224 X:C(Ind)[T.1] -1.8466 0.353 -5.239 0.000 -2.538 -1.155
==========
Omnibus: 4.301 Durbin-Watson: 1.985 Prob(Omnibus): 0.116 Jarque-Bera (JB): red 4.811 Skew: 0.065 Prob(JB): 0.0902 Kurtosis: 3.314 Cond. No. 119.
==========
Notes: [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

El término de interacción en el modelo es negativo (-1.8466) y es estadísticamente significativo, esto sugiere que la pendiente de la recta de regresión para X cuando Ind = 1 es menor en 1.8466 unidades que la pendiente de la recta de regresión para X cuando Ind = 0.

[breakable, size=fbox, boxrule=1pt, pad at breakable=1mm,colback=cellbackground, colframe=cellborder]: {},codes*=] intercept [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40. redparams[[rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Intercept[rgb]0.73,0.13,0.13']

```
slope˙X [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40.
    redparams[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’]
interaction [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40.
    redparams[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X:C(Ind)[T.
    red1][rgb]0.73,0.13,0.13’]
    x˙vals [rgb]0.40,0.40,0.40= np[rgb]0.40,0.40,0.40.
    redlinspace(data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’][rgb]0.40,0.40,0.40.
    redmin(), data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’][rgb]0.40,0.40,0.40.
    redmax(), [rgb]0.40,0.40,0.40100)
    y˙pred˙0 [rgb]0.40,0.40,0.40= intercept [rgb]0.40,0.40,0.40+
slope˙X [rgb]0.40,0.40,0.40* x˙vals y˙pred˙1 [rgb]0.40,0.40,0.40=
intercept [rgb]0.40,0.40,0.40+ model˙with˙interaction[rgb]0.40,0.40,0.40.
    redparams[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13C(Ind)[T.
    red1][rgb]0.73,0.13,0.13’] [rgb]0.40,0.40,0.40+ (slope˙X [rgb]0.40,0.40,0.40+
interaction) [rgb]0.40,0.40,0.40* x˙vals
    fig, axes [rgb]0.40,0.40,0.40= plt[rgb]0.40,0.40,0.40.
    redsubplots([rgb]0.40,0.40,0.401, [rgb]0.40,0.40,0.403,
figsize[rgb]0.40,0.40,0.40=([rgb]0.40,0.40,0.4015, [rgb]0.40,0.40,0.405))
    title˙interaction˙ind0 [rgb]0.40,0.40,0.40= [rgb]0.73,0.13,0.13r[rgb]0.73,0.13,0.13”[rgb]0.73,0.13,0.13($Ind
= 0$): $Y = [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47−0”[rgb]0.73,0.13,0.13
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47−X”[rgb]0.73,0.13,0.13X$[rgb]0.73,0.13,0.13
title˙interaction˙ind1 [rgb]0.40,0.40,0.40= [rgb]0.73,0.13,0.13r[rgb]0.73,0.13,0.13”[rgb]0.73,0.13,0.13($Ind
= 1$): $Y = ([rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47−0”[rgb]0.73,0.13,0.13
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.73,0.13,0.13−[rgb]0.73,0.13,0.13Ind=1″)
+ ([rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47−X”[rgb]0.73,0.13,0.13
+ [rgb]0.73,0.13,0.13“[rgb]0.73,0.13,0.13beta˙[rgb]0.64,0.35,0.47−interaction”[rgb]0.73,0.13,0.13)X$[rgb]0.
    axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[r
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
    red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
    red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
    red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.plot(x˙vals, y˙pred˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13black[rgb]0.73,0.13,0.13’,
    red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
    redset˙title(title˙interaction˙ind0) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
    redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
    axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[r
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
    red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
    red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
    red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
```

= 1[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.plot(x˙vals, y˙pred˙1,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13black[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙title(title˙interaction˙ind1) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[r
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 1[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.plot(x˙vals, y˙pred˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redplot(x˙vals, y˙pred˙1, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redset˙title([rgb]0.73,0.13,0.13”[rgb]0.73,0.13,0.13Ind = 0 y Ind
= 1[rgb]0.73,0.13,0.13”) axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.402][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
plt[rgb]0.40,0.40,0.40.tight˙layout() plt[rgb]0.40,0.40,0.40.show()

max size=0.90.9punto$_1$*f iles/punto*$_1$$_2$6$_0$*.png*

[breakable, size=fbox, boxrule=1pt, pad at
breakb*=1mm,colback=cellbackground, colframe=colframe]:
{},codes*=] fig, axes [rgb]0.40,0.40,0.40= plt[rgb]0.40,0.40,0.40.
redsubplots([rgb]0.40,0.40,0.401, [rgb]0.40,0.40,0.402,
figsize[rgb]0.40,0.40,0.40=([rgb]0.40,0.40,0.4015, [rgb]0.40,0.40,0.405))
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[r
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],

red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red2, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 1[rgb]0.73,0.13,0.13’, alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red2, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.plot(x˙vals, pred˙Y˙ind˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión Ind
= 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.plot(x˙vals,
pred˙Y˙ind˙1, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red linewidth[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.402,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
redset˙title([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Sin
interacción[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’)
axes[[rgb]0.40,0.40,0.400][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.scatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[r
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.400][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 0[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404,
red alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.
red2) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redscatter(data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’],
red data[data[[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind[rgb]0.73,0.13,0.13’]
[rgb]0.40,0.40,0.40== [rgb]0.40,0.40,0.401][[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,0.13’],
red color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Ind
= 1[rgb]0.73,0.13,0.13’, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404,
red alpha[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400.2)
axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.plot(x˙vals, y˙pred˙0,
color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’,
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión
Ind = 0[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.

redplot(x˙vals, y˙pred˙1, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13red[rgb]0.73,0.13,0
red label[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Regresión Ind = 1[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙title([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Con interacción[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.
redset˙xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13X[rgb]0.73,0.13,0.13’) axes[[rgb]0.40,0.40,0.401][rgb]0.40,0.40,0.40.set˙ylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Y[rgb]0.73,0.13,

[breakable, size=fbox, boxrule=.5pt, pad at break*=1mm, opacityfill=0.2]:
{},codes*=] Text(0, 0.5, ’Y’)

max size=0.90.9punto$_1$*f iles/punto*$_1$$_2$7$_1$*.png*

Los gráficos y el análisis del modelo indica que las interacciones son significativas. Al implementar el modelo con interaccieon,se observa que no solo se alinea mejor con los datos observados, sino que también proporciona un marco más completo y realista para entender y predecir la variable dependiente Y.

[breakable, size=fbox, boxrule=1pt, pad at break*=1mm, colback=cellbackground, colframe=cellborder]:
{},codes*=] residuals [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40.
redresid fitted [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40.
redfittedvalues
[rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Linealidad y Homoscedasticidad[rgb]0.73,0.13,0.13’, separator) plt[rgb]0.40,0.40,0.40.
redscatter(fitted, residuals, s[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.404, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13blue[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.axhline(y[rgb]0.40,0.40,0.40=[rgb]0.40,0.40,0.400, color[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13r[rgb]0.73,0.13,0.13’,
red linestyle[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13-- [rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.xlabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Valores Ajustados[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.
redylabel([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Residuales[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Residuales vs Valores Ajustados[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.show()
[rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Independencia[rgb]0.73,0.13,0.13’,
red separator) [rgb]0.24,0.48,0.48# *Durbin-Watson* dw [rgb]0.40,0.40,0.40= durbin˙watson(residuals) [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Estadístico de Durbin-Watson:[rgb]0.73,0.13,0.13’, dw)
[rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Normalidad[rgb]0.73,0.13,0.13’,
red separator) [rgb]0.24,0.48,0.48# *Gráfico Q-Q de los residuales* sm[rgb]0.40,0.40,0.40.qqplot(residuals, line[rgb]0.40,0.40,0.40=[rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13s[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.title([rgb]0.73,0.13,0.13’[rgb]0.73,0.13,0.13Gráfico Q-Q de los residuales[rgb]0.73,0.13,0.13’) plt[rgb]0.40,0.40,0.40.show()

[rgb]0.24,0.48,0.48# *Prueba Shapiro-Wilk de normalidad* [rgb]0.00,0.50,0.00print(stats[rgb]0.40,0.40,0.40.shapiro(residuals))
[rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13Multicolinealidad[rgb]0.73,0.13,0.13', red separator) [rgb]0.24,0.48,0.48# *Calculando el VIF para cada variable independiente* [rgb]0.00,0.50,0.00**from** [rgb]0.00,0.00,1.00**statsmodels**[rgb]0.00,0.00,1.00**.** red[rgb]0.00,0.00,1.00**stats**[rgb]0.00,0.00,1.00**.**[rgb]0.00,0.00,1.00**outliers˙influence** [rgb]0.00,0.50,0.00**import** variance˙inflation˙factor variables [rgb]0.40,0.40,0.40= model˙with˙interaction[rgb]0.40,0.40,0.40. redmodel[rgb]0.40,0.40,0.40.exog vif [rgb]0.40,0.40,0.40= [variance˙inflation˙factor(variables, i) [rgb]0.00,0.50,0.00**for** i [rgb]0.67,0.13,1.00**in** [rgb]0.00,0.50,0.00range(variables[rgb]0.40,0.40,0.40. redshape[[rgb]0.40,0.40,0.401])] [rgb]0.00,0.50,0.00print([rgb]0.73,0.13,0.13'[rgb]0.73,0.13,0.13VIFs: red[rgb]0.73,0.13,0.13', vif)
[commandchars= {},codes*=] Linealidad y Homoscedasticidad ==========

max size=0.90.9punto$_1$*f*$_i$*les/punto*$_1$$_2$9$_1$*.png*

[commandchars= {},codes*=] Independencia ==========
Estadístico de Durbin-Watson: 1.9848054259274441 Normalidad ==========

max size=0.90.9punto$_1$*f*$_i$*les/punto*$_1$$_2$9$_3$*.png*

[commandchars= {},codes*=] ShapiroResult(statistic=0.9961398243904114, red pvalue=0.013781944289803505) Multicolinealidad ========== VIFs: [8. red754810845175124, 18.852602858861175, 1.0760572580613532, 18. red983379216762742]