Regresiones tamaño por tipo

D. Daza

junio 13, 2023 | 11:01:45 | CDT

Table of Contents

[1 Modelos 2](#_Toc137550951)

[1.1 Cargas vs Tamaño y tipo FORMALES 2](#_Toc137550952)

[1.2 Cargas vs Tamaño y tipo INFORMALES 9](#_Toc137550953)

[2. Comprobación de supuestos 16](#_Toc137550954)

[2.1 Normalidad de residuos 16](#_Toc137550955)

[2.2 Independencia de los errores 17](#_Toc137550956)

[2.3 Linealidad de las relaciones entre las variables 18](#_Toc137550957)

[2.4 Homocedasticidad 21](#_Toc137550958)

# 1 Modelos

Las variables menos significativas en general son tasa de desocupación y densidad pob

## 1.1 Cargas vs Tamaño y tipo FORMALES

#############MICROEMPRESAS  
  
# RL micro formales simple  
mif\_reg1 <- lm(p\_mi\_f ~ ca\_xe, data = base)  
  
summary(mif\_reg1)

Call:  
lm(formula = p\_mi\_f ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.020307 -0.009399 -0.003038 0.008070 0.029376   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.170e-02 1.090e-02 2.909 0.00589 \*\*  
ca\_xe -2.785e-06 3.921e-06 -0.710 0.48155   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.01313 on 40 degrees of freedom  
Multiple R-squared: 0.01246, Adjusted R-squared: -0.01223   
F-statistic: 0.5047 on 1 and 40 DF, p-value: 0.4815

# RL micro formales controles  
mif\_reg2 <- lm(p\_mi\_f ~ ca\_xe +  
 delitos\_2019 +  
 p\_informal + pobreza +   
 razondeingreso,  
 data = base)  
  
summary(mif\_reg2)

Call:  
lm(formula = p\_mi\_f ~ ca\_xe + delitos\_2019 + p\_informal + pobreza +   
 razondeingreso, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.0064992 -0.0018606 -0.0004239 0.0013278 0.0091022   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 1.872e-01 9.245e-03 20.253 <2e-16 \*\*\*  
ca\_xe 1.298e-07 1.088e-06 0.119 0.9057   
delitos\_2019 -1.841e-07 9.015e-08 -2.042 0.0486 \*   
p\_informal -2.047e-01 1.235e-02 -16.566 <2e-16 \*\*\*  
pobreza 1.405e-04 7.146e-05 1.967 0.0570 .   
razondeingreso 6.091e-02 4.230e-02 1.440 0.1585   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.003349 on 36 degrees of freedom  
Multiple R-squared: 0.9422, Adjusted R-squared: 0.9342   
F-statistic: 117.4 on 5 and 36 DF, p-value: < 2.2e-16

vif(mif\_reg2)

ca\_xe delitos\_2019 p\_informal pobreza razondeingreso   
 1.185107 1.244115 2.767394 2.551206 1.461100

# Obtiene un data frame de los resultados del modelo  
mif2\_tidy <- broom::tidy(mif\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mif2\_table <- kable(mif2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y microempresas formales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mif2\_table

Cargas administrativas y microempresas formales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.1872481 | 0.0092454 | 20.2530208 | 0.0000000 |
| ca\_xe | 0.0000001 | 0.0000011 | 0.1193056 | 0.9056965 |
| delitos\_2019 | -0.0000002 | 0.0000001 | -2.0417400 | 0.0485573 |
| p\_informal | -0.2046639 | 0.0123547 | -16.5656174 | 0.0000000 |
| pobreza | 0.0001405 | 0.0000715 | 1.9667809 | 0.0569535 |
| razondeingreso | 0.0609070 | 0.0422983 | 1.4399382 | 0.1585251 |

#############PEQUEÑAS EMPRESAS  
  
# RL pequeñas formales simple  
pef\_reg1 <- lm(p\_pe\_f ~ ca\_xe, data = base)  
  
summary(pef\_reg1)

Call:  
lm(formula = p\_pe\_f ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.070308 -0.022883 0.001705 0.027579 0.078619   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 1.268e-01 2.910e-02 4.356 8.97e-05 \*\*\*  
ca\_xe 4.290e-06 1.047e-05 0.410 0.684   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.03507 on 40 degrees of freedom  
Multiple R-squared: 0.00418, Adjusted R-squared: -0.02072   
F-statistic: 0.1679 on 1 and 40 DF, p-value: 0.6842

# RL pequeñas formales controles  
pef\_reg2 <- lm(p\_pe\_f ~ ca\_xe +  
 escolaridadprom +  
 p\_informal + tasadesocupacion,  
 data = base)  
  
summary(pef\_reg2)

Call:  
lm(formula = p\_pe\_f ~ ca\_xe + escolaridadprom + p\_informal +   
 tasadesocupacion, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.037462 -0.015746 -0.001579 0.014261 0.052234   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 7.225e-01 8.566e-02 8.435 3.83e-10 \*\*\*  
ca\_xe -4.297e-07 7.060e-06 -0.061 0.951800   
escolaridadprom -1.721e-02 4.207e-03 -4.090 0.000224 \*\*\*  
p\_informal -5.476e-01 6.573e-02 -8.332 5.18e-10 \*\*\*  
tasadesocupacion 3.586e+00 9.460e-01 3.791 0.000537 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.02133 on 37 degrees of freedom  
Multiple R-squared: 0.6594, Adjusted R-squared: 0.6226   
F-statistic: 17.91 on 4 and 37 DF, p-value: 2.933e-08

vif(pef\_reg2)

ca\_xe escolaridadprom p\_informal tasadesocupacion   
 1.229815 1.437588 1.931876 1.778289

# Obtiene un data frame de los resultados del modelo  
pef2\_tidy <- broom::tidy(pef\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
pef2\_table <- kable(pef2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y pequeñas empresas formales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
pef2\_table

Cargas administrativas y pequeñas empresas formales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.7224997 | 0.0856566 | 8.4348404 | 0.0000000 |
| ca\_xe | -0.0000004 | 0.0000071 | -0.0608574 | 0.9518004 |
| escolaridadprom | -0.0172055 | 0.0042067 | -4.0899907 | 0.0002235 |
| p\_informal | -0.5476450 | 0.0657303 | -8.3316935 | 0.0000000 |
| tasadesocupacion | 3.5861353 | 0.9460031 | 3.7908281 | 0.0005369 |

#############MEDIANAS EMPRESAS  
  
# RL medianas formales simple  
mef\_reg1 <- lm(p\_me\_f ~ ca\_xe, data = base)  
  
summary(mef\_reg1)

Call:  
lm(formula = p\_me\_f ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.062059 -0.030981 0.004536 0.026080 0.061134   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.210e-01 2.724e-02 8.113 5.56e-10 \*\*\*  
ca\_xe 6.463e-06 9.799e-06 0.660 0.513   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.03283 on 40 degrees of freedom  
Multiple R-squared: 0.01076, Adjusted R-squared: -0.01397   
F-statistic: 0.4351 on 1 and 40 DF, p-value: 0.5133

# RL medianas formales controles  
mef\_reg2 <- lm(p\_me\_f ~ ca\_xe +  
 p\_informal + pobreza +   
 densidadpobkm2,  
 data = base)  
  
summary(mef\_reg2)

Call:  
lm(formula = p\_me\_f ~ ca\_xe + p\_informal + pobreza + densidadpobkm2,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.046170 -0.015144 0.000469 0.015581 0.054803   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.775e-01 7.510e-02 5.027 1.3e-05 \*\*\*  
ca\_xe 4.740e-06 8.182e-06 0.579 0.56585   
p\_informal -1.904e-01 1.060e-01 -1.797 0.08057 .   
pobreza 7.358e-04 6.121e-04 1.202 0.23697   
densidadpobkm2 -2.616e-06 8.349e-07 -3.133 0.00338 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.02672 on 37 degrees of freedom  
Multiple R-squared: 0.3938, Adjusted R-squared: 0.3283   
F-statistic: 6.01 on 4 and 37 DF, p-value: 0.0007876

vif(mef\_reg2)

ca\_xe p\_informal pobreza densidadpobkm2   
 1.052391 3.199185 2.941266 1.286474

# Obtiene un data frame de los resultados del modelo  
mef2\_tidy <- broom::tidy(mef\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mef2\_table <- kable(mef2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y medianas empresas formales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mef2\_table

Cargas administrativas y medianas empresas formales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.3775017 | 0.0750953 | 5.0269700 | 0.0000130 |
| ca\_xe | 0.0000047 | 0.0000082 | 0.5793653 | 0.5658538 |
| p\_informal | -0.1903783 | 0.1059674 | -1.7965735 | 0.0805713 |
| pobreza | 0.0007358 | 0.0006121 | 1.2020886 | 0.2369665 |
| densidadpobkm2 | -0.0000026 | 0.0000008 | -3.1327835 | 0.0033792 |

#############GRANDES EMPRESAS  
  
# RL grandes formales simple  
grf\_reg1 <- lm(p\_gr\_f ~ ca\_xe, data = base)  
  
summary(grf\_reg1)

Call:  
lm(formula = p\_gr\_f ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.087985 -0.034054 -0.004672 0.033833 0.089913   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.598e-01 3.881e-02 6.694 5.02e-08 \*\*\*  
ca\_xe -9.479e-06 1.396e-05 -0.679 0.501   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04678 on 40 degrees of freedom  
Multiple R-squared: 0.01139, Adjusted R-squared: -0.01332   
F-statistic: 0.4609 on 1 and 40 DF, p-value: 0.5011

# RL grandes formales controles  
grf\_reg2 <- lm(p\_gr\_f ~ ca\_xe +  
 escolaridadprom +  
 densidadpobkm2,  
 data = base)  
  
summary(grf\_reg2)

Call:  
lm(formula = p\_gr\_f ~ ca\_xe + escolaridadprom + densidadpobkm2,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.093893 -0.027783 0.003919 0.028987 0.068660   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 4.877e-01 8.811e-02 5.535 2.47e-06 \*\*\*  
ca\_xe -1.624e-05 1.191e-05 -1.364 0.1806   
escolaridadprom -1.797e-02 7.366e-03 -2.440 0.0195 \*   
densidadpobkm2 -2.431e-06 1.231e-06 -1.975 0.0555 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.03953 on 38 degrees of freedom  
Multiple R-squared: 0.3295, Adjusted R-squared: 0.2765   
F-statistic: 6.224 on 3 and 38 DF, p-value: 0.00152

vif(grf\_reg2)

ca\_xe escolaridadprom densidadpobkm2   
 1.018314 1.283569 1.277850

# Obtiene un data frame de los resultados del modelo  
grf2\_tidy <- broom::tidy(grf\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
grf2\_table <- kable(grf2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y grandes empresas formales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
grf2\_table

Cargas administrativas y grandes empresas formales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.4877154 | 0.0881109 | 5.535242 | 0.0000025 |
| ca\_xe | -0.0000162 | 0.0000119 | -1.363896 | 0.1806253 |
| escolaridadprom | -0.0179727 | 0.0073664 | -2.439811 | 0.0194717 |
| densidadpobkm2 | -0.0000024 | 0.0000012 | -1.975384 | 0.0555207 |

## 1.2 Cargas vs Tamaño y tipo INFORMALES

#############MICROEMPRESAS  
  
# RL micro informales simple  
mii\_reg1 <- lm(p\_mi\_i ~ ca\_xe, data = base)  
  
summary(mii\_reg1)

Call:  
lm(formula = p\_mi\_i ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.098749 -0.035497 0.003238 0.034587 0.068382   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 8.945e-01 3.792e-02 23.59 <2e-16 \*\*\*  
ca\_xe 9.419e-06 1.364e-05 0.69 0.494   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04571 on 40 degrees of freedom  
Multiple R-squared: 0.01178, Adjusted R-squared: -0.01293   
F-statistic: 0.4767 on 1 and 40 DF, p-value: 0.4939

# RL micro informales controles  
mii\_reg2 <- lm(p\_mi\_i ~ ca\_xe +  
 escolaridadprom +  
 p\_formal + pobreza,  
 data = base)  
  
summary(mii\_reg2)

Call:  
lm(formula = p\_mi\_i ~ ca\_xe + escolaridadprom + p\_formal + pobreza,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.0257085 -0.0072793 -0.0008185 0.0064796 0.0213409   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 9.183e-01 3.149e-02 29.159 < 2e-16 \*\*\*  
ca\_xe -4.383e-06 3.345e-06 -1.310 0.19811   
escolaridadprom 8.032e-03 2.261e-03 3.552 0.00106 \*\*   
p\_formal -2.024e+00 1.208e-01 -16.747 < 2e-16 \*\*\*  
pobreza 6.083e-04 2.385e-04 2.550 0.01503 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.01101 on 37 degrees of freedom  
Multiple R-squared: 0.9469, Adjusted R-squared: 0.9412   
F-statistic: 165 on 4 and 37 DF, p-value: < 2.2e-16

vif(mii\_reg2)

ca\_xe escolaridadprom p\_formal pobreza   
 1.035041 1.557342 2.167447 2.628832

# Obtiene un data frame de los resultados del modelo  
mii2\_tidy <- broom::tidy(mii\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mii2\_table <- kable(mii2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y microempresas informales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mii2\_table

Cargas administrativas y microempresas informales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.9182946 | 0.0314927 | 29.158964 | 0.0000000 |
| ca\_xe | -0.0000044 | 0.0000033 | -1.310488 | 0.1981076 |
| escolaridadprom | 0.0080321 | 0.0022610 | 3.552369 | 0.0010621 |
| p\_formal | -2.0235634 | 0.1208334 | -16.746719 | 0.0000000 |
| pobreza | 0.0006083 | 0.0002385 | 2.550444 | 0.0150303 |

#############PEQUEÑAS EMPRESAS  
  
# RL pequeñas informales simple  
pei\_reg1 <- lm(p\_pe\_i ~ ca\_xe, data = base)  
  
summary(pei\_reg1)

Call:  
lm(formula = p\_pe\_i ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.18540 -0.08646 -0.02111 0.08592 0.22724   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 5.445e-01 9.330e-02 5.835 8.05e-07 \*\*\*  
ca\_xe -7.936e-06 3.357e-05 -0.236 0.814   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.1125 on 40 degrees of freedom  
Multiple R-squared: 0.001396, Adjusted R-squared: -0.02357   
F-statistic: 0.0559 on 1 and 40 DF, p-value: 0.8143

# RL pequeñas informales controles  
pei\_reg2 <- lm(p\_pe\_i ~ ca\_xe +  
 escolaridadprom +  
 p\_formal + tasadesocupacion + pobreza +   
 densidadpobkm2,  
 data = base)  
  
summary(pei\_reg2)

Call:  
lm(formula = p\_pe\_i ~ ca\_xe + escolaridadprom + p\_formal + tasadesocupacion +   
 pobreza + densidadpobkm2, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.073188 -0.020724 0.002904 0.020822 0.064412   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.415e-01 1.115e-01 2.166 0.037243 \*   
ca\_xe -1.736e-05 1.183e-05 -1.468 0.151145   
escolaridadprom 4.329e-02 1.016e-02 4.260 0.000146 \*\*\*  
p\_formal -4.082e+00 5.025e-01 -8.123 1.45e-09 \*\*\*  
tasadesocupacion -4.182e+00 1.577e+00 -2.651 0.011968 \*   
pobreza 3.313e-03 7.802e-04 4.246 0.000152 \*\*\*  
densidadpobkm2 5.100e-06 1.421e-06 3.588 0.001008 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.03562 on 35 degrees of freedom  
Multiple R-squared: 0.9123, Adjusted R-squared: 0.8973   
F-statistic: 60.7 on 6 and 35 DF, p-value: < 2.2e-16

vif(pei\_reg2)

ca\_xe escolaridadprom p\_formal tasadesocupacion   
 1.237335 3.006729 3.582748 1.772568   
 pobreza densidadpobkm2   
 2.688761 2.097435

# Obtiene un data frame de los resultados del modelo  
pei2\_tidy <- broom::tidy(pei\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
pei2\_table <- kable(pei2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y pequeñas empresas informales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
pei2\_table

Cargas administrativas y pequeñas empresas informales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.2415406 | 0.1115369 | 2.165567 | 0.0372433 |
| ca\_xe | -0.0000174 | 0.0000118 | -1.467589 | 0.1511455 |
| escolaridadprom | 0.0432871 | 0.0101611 | 4.260076 | 0.0001462 |
| p\_formal | -4.0815067 | 0.5024558 | -8.123116 | 0.0000000 |
| tasadesocupacion | -4.1820963 | 1.5774676 | -2.651146 | 0.0119679 |
| pobreza | 0.0033126 | 0.0007802 | 4.245906 | 0.0001524 |
| densidadpobkm2 | 0.0000051 | 0.0000014 | 3.588418 | 0.0010077 |

#############MEDIANAS EMPRESAS  
  
# RL medianas informales simple  
mei\_reg1 <- lm(p\_me\_i ~ ca\_xe, data = base)  
  
summary(mei\_reg1)

Call:  
lm(formula = p\_me\_i ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.113497 -0.053721 -0.001924 0.045187 0.134523   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.482e-01 5.279e-02 6.596 6.88e-08 \*\*\*  
ca\_xe 1.229e-05 1.899e-05 0.647 0.521   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.06362 on 40 degrees of freedom  
Multiple R-squared: 0.01036, Adjusted R-squared: -0.01438   
F-statistic: 0.4188 on 1 and 40 DF, p-value: 0.5213

# RL medianas informales controles  
mei\_reg2 <- lm(p\_me\_i ~ ca\_xe +  
 escolaridadprom +  
 p\_formal,  
 data = base)  
  
summary(mei\_reg2)

Call:  
lm(formula = p\_me\_i ~ ca\_xe + escolaridadprom + p\_formal, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.086975 -0.033926 -0.000843 0.034559 0.069651   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.619e-01 9.319e-02 2.810 0.00778 \*\*   
ca\_xe 4.758e-07 1.315e-05 0.036 0.97133   
escolaridadprom 2.145e-02 8.096e-03 2.649 0.01168 \*   
p\_formal -2.545e+00 3.689e-01 -6.899 3.36e-08 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04341 on 38 degrees of freedom  
Multiple R-squared: 0.5624, Adjusted R-squared: 0.5278   
F-statistic: 16.28 on 3 and 38 DF, p-value: 5.81e-07

vif(mei\_reg2)

ca\_xe escolaridadprom p\_formal   
 1.030267 1.285430 1.300421

# Obtiene un data frame de los resultados del modelo  
mei2\_tidy <- broom::tidy(mei\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mei2\_table <- kable(mei2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y medianas empresas informales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mei2\_table

Cargas administrativas y medianas empresas informales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.2618727 | 0.0931899 | 2.8100980 | 0.0077847 |
| ca\_xe | 0.0000005 | 0.0000132 | 0.0361791 | 0.9713289 |
| escolaridadprom | 0.0214486 | 0.0080955 | 2.6494523 | 0.0116823 |
| p\_formal | -2.5449171 | 0.3688559 | -6.8994890 | 0.0000000 |

#############GRANDES EMPRESAS  
  
# RL grandes informales simple  
gri\_reg1 <- lm(p\_gr\_i ~ ca\_xe, data = base)  
  
summary(gri\_reg1)

Call:  
lm(formula = p\_gr\_i ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.120685 -0.027061 0.001547 0.036332 0.153241   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.035e-01 5.100e-02 3.990 0.000274 \*\*\*  
ca\_xe 5.283e-05 1.835e-05 2.879 0.006370 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.06146 on 40 degrees of freedom  
Multiple R-squared: 0.1717, Adjusted R-squared: 0.151   
F-statistic: 8.291 on 1 and 40 DF, p-value: 0.00637

# RL grandes informales controles  
gri\_reg2 <- lm(p\_gr\_i ~ ca\_xe +  
 delitos\_2019 +  
 densidadpobkm2,  
 data = base)  
  
summary(gri\_reg2)

Call:  
lm(formula = p\_gr\_i ~ ca\_xe + delitos\_2019 + densidadpobkm2,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.096479 -0.036934 -0.003167 0.030282 0.110889   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.742e-01 5.175e-02 5.299 5.2e-06 \*\*\*  
ca\_xe 3.621e-05 1.646e-05 2.200 0.033933 \*   
delitos\_2019 -4.858e-06 1.344e-06 -3.613 0.000873 \*\*\*  
densidadpobkm2 4.289e-06 1.442e-06 2.975 0.005064 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.05145 on 38 degrees of freedom  
Multiple R-squared: 0.4487, Adjusted R-squared: 0.4051   
F-statistic: 10.31 on 3 and 38 DF, p-value: 4.225e-05

vif(gri\_reg2)

ca\_xe delitos\_2019 densidadpobkm2   
 1.148225 1.172879 1.034626

# Obtiene un data frame de los resultados del modelo  
gri2\_tidy <- broom::tidy(gri\_reg2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
gri2\_table <- kable(gri2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y grandes empresas informales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
gri2\_table

Cargas administrativas y grandes empresas informales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.2742118 | 0.0517491 | 5.298868 | 0.0000052 |
| ca\_xe | 0.0000362 | 0.0000165 | 2.200346 | 0.0339333 |
| delitos\_2019 | -0.0000049 | 0.0000013 | -3.613160 | 0.0008731 |
| densidadpobkm2 | 0.0000043 | 0.0000014 | 2.975474 | 0.0050639 |

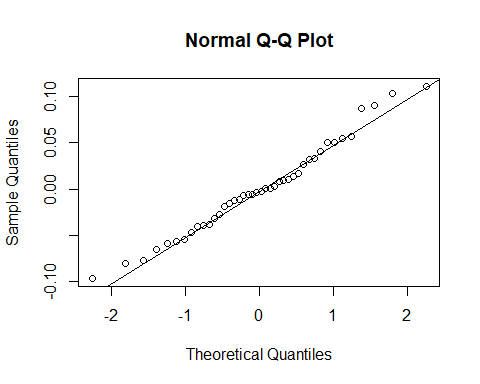
# 2. Comprobación de supuestos

gri\_reg2

## 2.1 Normalidad de residuos

Normalidad de los residuos: Puedes usar un gráfico Q-Q de los residuos para verificar la normalidad. Los puntos deben caer aproximadamente en una línea recta si los residuos están normalmente distribuidos.

# Generar residuos  
residuos <- residuals(gri\_reg2)  
  
# Crear gráfico Q-Q  
qqnorm(residuos)  
qqline(residuos)



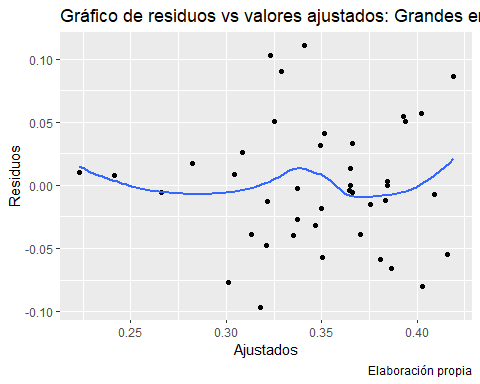
#Para usar la prueba de Shapiro-Wilk para verificar la normalidad, generalmente mirarías el valor p. Si el valor p es menor que 0.05, entonces tienes evidencia para rechazar la hipótesis nula y puedes concluir que tus datos probablemente no provienen de una distribución normal. Si el valor p es mayor que 0.05, entonces no tienes suficiente evidencia para rechazar la hipótesis nula y puedes concluir que tus datos pueden provenir de una distribución normal.  
  
shapiro.test(residuos)

Shapiro-Wilk normality test  
  
data: residuos  
W = 0.97906, p-value = 0.6253

## 2.2 Independencia de los errores

Un gráfico de los residuos en función de los valores ajustados o el orden temporal puede ayudar a identificar si existe una correlación entre los errores.

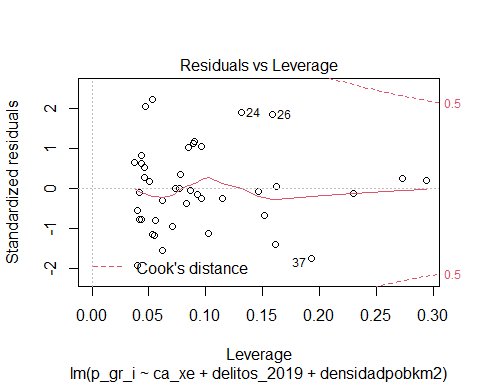
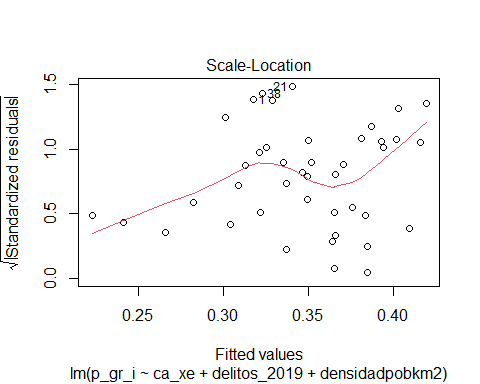
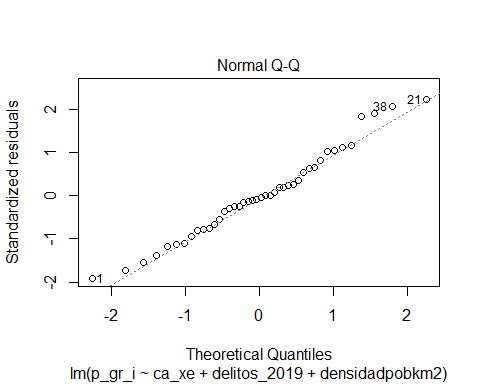
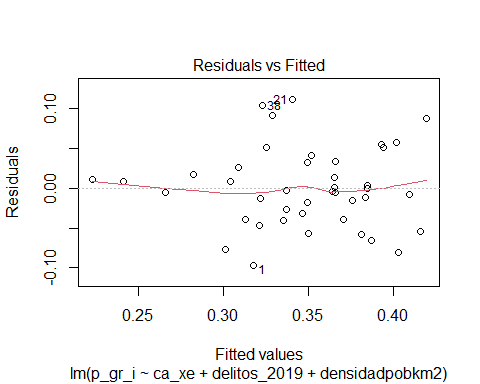
# Crear un gráfico de residuos vs valores ajustados  
ggplot(data = data.frame(Residuos = residuos, Ajustados = fitted(gri\_reg2)), aes(x = Ajustados, y = Residuos)) +  
 geom\_point() +  
 geom\_smooth(se = FALSE) +  
 labs(title = "Gráfico de residuos vs valores ajustados: Grandes empresas informales",  
 caption = "Elaboración propia")



## 2.3 Linealidad de las relaciones entre las variables

Para verificar la linealidad, puedes hacer gráficos de residuos contra cada variable predictora.

# Gráfico de residuos contra cada variable predictora  
  
plot(gri\_reg2)



## 2.4 Homocedasticidad

Para comprobar la homocedasticidad (varianza constante de los errores), puedes hacer un gráfico de residuos contra valores ajustados. La dispersión de los puntos debe ser más o menos constante a lo largo del rango de valores ajustados.

#También puedes usar la prueba de Breusch-Pagan para comprobar la homocedasticidad. Si el valor p es significativo (por lo general, p < 0.05), entonces tienes evidencia de heterocedasticidad.  
  
library(lmtest)  
  
# Prueba de Breusch-Pagan  
bptest(gri\_reg2)

studentized Breusch-Pagan test  
  
data: gri\_reg2  
BP = 0.90505, df = 3, p-value = 0.8242