

EXAMEN 1

En la siguiente tabla aparecen algunos de los datos, y las variables que serán objeto de estudio, siendo "luz" la variable dependiente y x0, x1, x2, x3, x4, x5 y mes las variables independientes.

Se llama la base de datos con el siguiente código = **Base1<-read.csv("examen1.csv")**

Y se instalan los siguientes paquetes

install.packages("plm")

install.packages("ggplot2")

library("plm")

library("ggplot2")

	clave	t	cve_ent	cve_mun	nom_mun	luz	x0	x1	x2	x3	x4	x5	mes
1	1001	214	1	1	Aguascalientes	46984964	26.94436	17754	5918	554457.41	13625	6.36342140	1
2	1001	215	1	1	Aguascalientes	54418648	28.83068	54418648	759	579727.04	13625	3.27513600	1
3	1001	216	1	1	Aguascalientes	52091332	26.45043	52091332	1289	590379.68	13625	5.84301570	1
4	1001	217	1	1	Aguascalientes	77462424	34.16601	77462424	3034	588190.02	13625	2.93892790	1
5	1001	218	1	1	Aguascalientes	61593976	36.36248	61593976	3179	573927.57	13625	1.12517440	1
6	1001	219	1	1	Aguascalientes	48427296	33.69948	48427296	5332	548776.12	13625	1.16465720	1
7	1001	220	1	1	Aguascalientes	43905296	26.62954	43905296	747	512853.87	13625	1.79602850	1
8	1001	221	1	1	Aguascalientes	39072976	28.16003	39072976	3	470421.66	13625	3.85937120	1
9	1001	222	1	1	Aguascalientes	36778764	28.98830	36778764	0	422031.94	13625	1.80774890	1
10	1001	223	1	1	Aguascalientes	23604476	29.28346	23604476	5697	369381.96	13625	4.54271550	1
11	1001	224	1	1	Aguascalientes	27699370	30.11087	27699370	3899	314032.92	13625	2.42655780	1
12	1001	225	1	1	Aguascalientes	25026574	30.65534	25026574	550	254118.91	13625	5.08338930	1
13	1001	226	1	1	Aguascalientes	22022192	30.27448	22022192	5575	197964.57	13625	2.55267520	1
14	1001	227	1	1	Aguascalientes	29780670	34.84234	29780670	3056	144576.72	13625	1.82203150	1
15	1001	228	1	1	Aguascalientes	33007834	34.49759	33007834	0	96107.65	13625	3.77686640	1
16	1001	229	1	1	Aguascalientes	16191345	28.61611	16191345	5227	55037.25	13625	2.65218590	1
17	1001	230	1	1	Aguascalientes	29029426	28.13513	29029426	2412	23967.83	13625	4.30587820	1
18	1001	231	1	1	Aguascalientes	14242743	33.68840	14242743	5394	4260.96	13625	4.20505620	1
19	1001	232	1	1	Aguascalientes	10445644	29.62430	10445644	5918	1479.41	13625	0.93273187	1

Tabla 1. Extracto de la base de datos

1. Incorporar la base de datos (panel) en R, y declararla como panel

Se utiliza el siguiente código para declarar la base de datos previamente incorporada a R, como panel =

Panel1 <- **pdata.frame**(Base1, index = c("clave", "t")) , y se utiliza el código **Panel1**, para extraer una parte de él.

	clave	t	cve_ent	cve_mun	nom_mun	luz	x0	x1	x2	x3	x4	x5	mes
1001-214	1001	214	1	1	Aguascalientes	46984964	26.94436	17754	5918	554457.41	13625	6.36342140	1
1001-215	1001	215	1	1	Aguascalientes	54418648	28.83068	54418648	759	579727.04	13625	3.27513600	1
1001-216	1001	216	1	1	Aguascalientes	52091332	26.45043	52091332	1289	590379.68	13625	5.84301570	1
1001-217	1001	217	1	1	Aguascalientes	77462424	34.16601	77462424	3034	588190.02	13625	2.93892790	1
1001-218	1001	218	1	1	Aguascalientes	61593976	36.36248	61593976	3179	573927.57	13625	1.12517440	1
1001-219	1001	219	1	1	Aguascalientes	48427296	33.69948	48427296	5332	548776.12	13625	1.16465720	1
1001-220	1001	220	1	1	Aguascalientes	43905296	26.62954	43905296	747	512853.87	13625	1.79602850	1
1001-221	1001	221	1	1	Aguascalientes	39072976	28.16003	39072976	3	470421.66	13625	3.85937120	1
1001-222	1001	222	1	1	Aguascalientes	36778764	28.98830	36778764	0	422031.94	13625	1.80774890	1
1001-223	1001	223	1	1	Aguascalientes	23604476	29.28346	23604476	5697	369381.96	13625	4.54271550	1
1001-224	1001	224	1	1	Aguascalientes	27699370	30.11087	27699370	3899	314032.92	13625	2.42655780	1
1001-225	1001	225	1	1	Aguascalientes	25026574	30.65534	25026574	550	254118.91	13625	5.08338930	1
1001-226	1001	226	1	1	Aguascalientes	22022192	30.27448	22022192	5575	197964.57	13625	2.55267520	1
1001-227	1001	227	1	1	Aguascalientes	29780670	34.84234	29780670	3056	144576.72	13625	1.82203150	1
1001-228	1001	228	1	1	Aguascalientes	33007834	34.49759	33007834	0	96107.65	13625	3.77686640	1
1001-229	1001	229	1	1	Aguascalientes	16191345	28.61611	16191345	5227	55037.25	13625	2.65218590	1
1001-230	1001	230	1	1	Aguascalientes	29029426	28.13513	29029426	2412	23967.83	13625	4.30587820	1
1001-231	1001	231	1	1	Aguascalientes	14242743	33.68840	14242743	5394	4260.96	13625	4.20505620	1
1001-232	1001	232	1	1	Aguascalientes	10445644	29.62430	10445644	5918	1479.41	13625	0.93273187	1
1001-233	1001	233	1	1	Aguascalientes	19074970	35.74582	19074970	5477	14203.11	13625	3.01948740	1
1001-234	1001	234	1	1	Aguascalientes	29589354	28.92847	29589354	1871	42120.59	13625	6.77180390	1
1001-235	1001	235	1	1	Aguascalientes	36703776	32.39971	36703776	0	83680.48	13625	6.92305520	1
1001-236	1001	236	1	1	Aguascalientes	36594636	29.15608	36594636	995	141736.99	13625	4.57367660	1
1001-237	1001	237	1	1	Aguascalientes	33110038	28.78511	33110038	4333	202393.61	13625	1.58119030	1
1001-238	1001	238	1	1	Aguascalientes	35645656	37.04787	35645656	3910	266797.66	13625	1.62941430	1
1001-239	1001	239	1	1	Aguascalientes	17165128	32.10748	17165128	2211	331644.69	13625	0.91822559	1
1001-240	1001	240	1	1	Aguascalientes	28080352	28.00923	28080352	5883	393842.75	13625	3.43637040	1
1001-241	1001	241	1	1	Aguascalientes	12591251	30.85542	12591251	5780	454201.25	13625	5.05498030	1
1001-242	1001	242	1	1	Aguascalientes	35416960	34.78976	35416960	0	502319.83	13625	2.45792790	1
1001-243	1001	243	1	1	Aguascalientes	37547384	29.02088	37547384	1363	541023.54	13625	5.97383980	1
1001-244	1001	244	1	1	Aguascalientes	50894644	25.39442	50894644	4073	569074.86	13625	5.29036860	1
1001-245	1001	245	1	1	Aguascalientes	93762648	33.17934	93762648	1996	585645.16	13625	0.26869878	2
1001-246	1001	246	1	1	Aguascalientes	59439880	28.69080	59439880	8	590557.19	13625	6.49012660	2
1001-247	1001	247	1	1	Aguascalientes	57696656	26.00605	57696656	2724	583573.93	13625	5.92888880	2
1001-248	1001	248	1	1	Aguascalientes	46855164	30.77989	46855164	0	565819.90	13625	1.00254080	2
1001-249	1001	249	1	1	Aguascalientes	40765602	27.38537	40765602	1380	538301.26	13625	6.58070700	2

2. Realizar una tabla de estadísticas descriptivas de las variables

Se realiza la siguiente regresión para conocer y presentar en una tabla de estadística descriptiva de las variables

reg1<-**lm**(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1)

summary(reg1)

```

Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-25833941  -219155  -115479   103442 255301830

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.450e+05  7.623e+04   7.149  8.8e-13 ***
x0          -4.200e+03  2.389e+03  -1.758   0.0788 .
x1           9.651e-01  4.350e-04 2218.670 < 2e-16 ***
x2           9.415e+01  2.012e+00  46.783 < 2e-16 ***
x3           1.528e+00  1.794e-02  85.183 < 2e-16 ***
x4          -8.289e+00  2.530e-01 -32.766 < 2e-16 ***
x5          -1.681e+03  3.550e+03  -0.473   0.6359
mes          -2.683e+05  1.435e+04 -18.690 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2775000 on 149930 degrees of freedom
Multiple R-squared:  0.9871,    Adjusted R-squared:  0.9871
F-statistic: 1.638e+06 on 7 and 149930 DF,  p-value: < 2.2e-16

```

Tabla 2. Estadísticos descriptivos

Esta tabla nos muestra que el intercepto, al igual que las variables x1, x2, x3, x4 y mes son significativas a más del 99%. Y las variables explican en un 0,9871 el modelo.

head(Base1)

```

> head(Base1)
  clave  t cve_ent cve_mun  nom_mun  luz  x0  x1  x2  x3  x4  x5 mes
1 1001 214      1      1 Aguascalientes 46984964 26.94436 17754 5918 554457.4 13625 6.363421 1
2 1001 215      1      1 Aguascalientes 54418648 28.83068 54418648 759 579727.0 13625 3.275136 1
3 1001 216      1      1 Aguascalientes 52091332 26.45043 52091332 1289 590379.7 13625 5.843016 1
4 1001 217      1      1 Aguascalientes 77462424 34.16601 77462424 3034 588190.0 13625 2.938928 1
5 1001 218      1      1 Aguascalientes 61593976 36.36248 61593976 3179 573927.6 13625 1.125174 1
6 1001 219      1      1 Aguascalientes 48427296 33.69948 48427296 5332 548776.1 13625 1.164657 1

```

3. Realizar regresión pooled-OLS

```
OLS<- lm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1)
```

```
summary(OLS)
```

```

Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-25833941  -219155  -115479   103442  255301830

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.450e+05  7.623e+04   7.149  8.8e-13 ***
x0          -4.200e+03  2.389e+03  -1.758  0.0788 .
x1           9.651e-01  4.350e-04 2218.670 < 2e-16 ***
x2           9.415e+01  2.012e+00  46.783 < 2e-16 ***
x3           1.528e+00  1.794e-02  85.183 < 2e-16 ***
x4          -8.289e+00  2.530e-01  -32.766 < 2e-16 ***
x5          -1.681e+03  3.550e+03  -0.473  0.6359
mes          -2.683e+05  1.435e+04 -18.690 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2775000 on 149930 degrees of freedom
Multiple R-squared:  0.9871,    Adjusted R-squared:  0.9871
F-statistic: 1.638e+06 on 7 and 149930 DF,  p-value: < 2.2e-16

```

Tabla 3. Regresion pooled OLS

Se observa que los coeficientes del intercepto, x0, x2, x3, x4 y mes son significativamente distintos de cero. Puesto que las estimaciones de los coeficientes x0 y x4 son negativas, cuando estas variables aumentan disminuye los niveles de luz. Por otro lado, puesto que las estimaciones de x1, x2, x3 y x5 son positivas, cuando éstas variables aumentan también lo hacen los niveles de luz. Finalmente, puesto que el p-valor es menor que 0.05, se tiene que el modelo es válido conjuntamente.

4. Determinar si la regresión es de efectos fijos o aleatorios

Efectos fijos

```
ef<-plm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1,index=c("clave", "t"),model="within")
```

summary(ef)

```

Call:
plm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + x5 + mes, data = Panel1,
     model = "within", index = c("clave", "t"))

Balanced Panel: n = 2458, T = 61, N = 149938

Residuals:
    Min.    1st Qu.    Median      Mean   3rd Qu.    Max.
-33171067  -227360   -105012      0    105518  248206726

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)
x0   -3.6719e+03    2.4037e+03  -1.5276  0.1266
x1    9.5833e-01    4.5622e-04 2100.5821 <2e-16 ***
x2    1.4429e+02    2.4767e+00  58.2594 <2e-16 ***
x3    2.1225e+00    2.1955e-02  96.6764 <2e-16 ***
x5   -2.5876e+03    3.5679e+03  -0.7253  0.4683
mes  -2.4835e+05    1.4323e+04 -17.3401 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    6.1184e+19
Residual Sum of Squares: 1.1296e+18
R-Squared:    0.98154
Adj. R-Squared: 0.98123
F-statistic: 1306690 on 6 and 147474 DF, p-value: < 2.22e-16

```


Tabla 4. Efectos fijos

La variable independiente no aparece ya que el modelo estimado es el de desviaciones con respecto a la media. De hecho, elimina todas las variables constantes en cada individuo, y en este caso elimina x4 porque presenta dicha característica.

Se presentan cambios además en las estimaciones de los coeficientes que son significativamente cercanos a 0, ya que son x1, x2, x3 y mes los que presentan esta condiciones. Además, las variables x0, x5 y mes al ser negativas cuando presentan un aumento hacen disminuir los niveles de luz, mientras que el incremento en las otras variables tiene el mismo efecto en los niveles de luz.

Efectos aleatorios

ea <- plm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1, index=c("clave", "t"), model="random")

summary(ea)

```
Call:
plm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + x5 + mes, data = Panel1,
     model = "random", index = c("clave", "t"))

Balanced Panel: n = 2458, T = 61, N = 149938

Effects:
              var   std.dev share
idiosyncratic 7.660e+12 2.768e+06    1
individual    0.000e+00 0.000e+00    0
theta: 0

Residuals:
      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
-25833941 -219155   -115479      0      103442 255301830

Coefficients:
              Estimate Std. Error z-value Pr(>|z|)
(Intercept)  5.4497e+05  7.6233e+04   7.1487 8.761e-13 ***
x0           -4.1998e+03  2.3895e+03  -1.7576  0.07881 .
x1            9.6510e-01  4.3499e-04 2218.6702 < 2.2e-16 ***
x2            9.4149e+01  2.0125e+00  46.7829 < 2.2e-16 ***
x3            1.5278e+00  1.7935e-02  85.1828 < 2.2e-16 ***
x4           -8.2892e+00  2.5298e-01 -32.7658 < 2.2e-16 ***
x5           -1.6808e+03  3.5498e+03  -0.4735  0.63586
mes          -2.6828e+05  1.4354e+04 -18.6898 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 8.9477e+19
Residual Sum of Squares: 1.1546e+18
R-Squared: 0.9871
Adj. R-Squared: 0.9871
ChiSq: 11469500 on 7 DF, p-value: < 2.22e-16
```

Tabla 5. Efectos aleatorios

Al observar el valor estimado de Theta se puede concluir que los efectos aleatorios se asemejan más a los resultados obtenidos por el modelo agrupado, ya que este es equivalente a 0.

Prueba para elegir efectos fijos o efectos aleatorios

phptest(ea)

Hausman Test

```
data: luz ~ x0 + x1 + x2 + x3 + x4 + x5 + mes
chisq = 2803.9, df = 6, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent
```

Tabla 6. Prueba Hausman Test

Al utilizar la prueba de Hausman en donde la hipótesis nula es que “el modelo de efectos aleatorios es mejor que el de efectos fijos”, y en este caso se rechaza dicha hipótesis ya que el P value es menor a 0,05, por lo tanto se prefieren los efectos fijos.

5. Probar si la base de datos tiene heteroscedasticidad, si la tiene, corregirla

```
library(lmtest)
```

```
bptest(ef)
```

```

      .      .      .
      .      .      .
      .      .      .

studentized Breusch-Pagan test

data: ef
BP = 9122, df = 7, p-value < 2.2e-16
```

Tabla 7. Prueba de heterocedasticidad

Se comprueba que la muestra si tiene heterocedasticidad , por lo tanto se debe corregir

```
vc_ef <- vcovHC(ef, type="HC0")
```

```
summary(vc_ef)
```

```

> vc_ef <- vcovHC(ef, type="HC0")
> summary(vc_ef)
      x0      x1      x2      x3      x5      mes
Min.   :-623063 Min.   :-21.362795 Min.   : -9313.82 Min.   : -71.4255 Min.   : -369731 Min.   : -623063
1st Qu.: -279626 1st Qu.: -1.540538 1st Qu.:   0.37 1st Qu.:  0.0098 1st Qu.:   27 1st Qu.:   175
Median :  -4693 Median : -0.034562 Median :  325.91 Median :  0.8721 Median :  7624 Median :  79315
Mean   : 729939 Mean   : -3.821632 Mean   : 27389.68 Mean   : 134.6943 Mean   : 2733578 Mean   : 23013770
3rd Qu.:  -17 3rd Qu.: -0.000076 3rd Qu.: 11512.18 3rd Qu.:  86.8448 3rd Qu.: 4543466 3rd Qu.: 4579150
Max.   :5381809 Max.   :  0.533241 Max.   :157867.29 Max.   :762.6223 Max.   :10703044 Max.   :132494165
```

Tabla 8. Corrección de heterocedasticidad

6. ¿Qué otras maneras de presentar las variables se te ocurren para que la Adj.R2 aumente de valor? ¿Es valida esa especificación?

Opción 1. Haciendo logarítmicas algunas variables

```
Panel1$x0log<-log(Panel1$x0)
```

```
Panel1$x4log<-log(Panel1$x4)
```

```
Panel1$x5log<-log(Panel1$x5)
```

```
Reglogx0<-lm(luz~x0log+x1+x2+x3+x4+x5+mes,data=Panel1)
```

summary(Reglogx0)

```
Call:
lm(formula = luz ~ x0log + x1 + x2 + x3 + x4 + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-25833797  -219114  -115996   103608  255302589

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.245e+05  2.417e+05   3.411 0.000647 ***
x0log       -1.194e+05  7.076e+04  -1.687 0.091530 .
x1           9.651e-01  4.350e-04  2218.669 < 2e-16 ***
x2           9.415e+01  2.012e+00   46.784 < 2e-16 ***
x3           1.528e+00  1.794e-02   85.182 < 2e-16 ***
x4          -8.289e+00  2.530e-01  -32.766 < 2e-16 ***
x5          -1.682e+03  3.550e+03  -0.474 0.635560
mes         -2.683e+05  1.435e+04  -18.689 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2775000 on 149930 degrees of freedom
Multiple R-squared:  0.9871,    Adjusted R-squared:  0.9871
F-statistic: 1.638e+06 on 7 and 149930 DF, p-value: < 2.2e-16
```

Tabla 9. Prueba 1

No aumento R2

Reglogx4<-lm(luz~x0+x1+x2+x3+x4log+x5+mes,data=Panel1)

summary(Reglogx4)

```
Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x4log + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-30594548  -241213  -99721   69812  257756369

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.480e+05  8.638e+04  10.974 <2e-16 ***
x0          -4.289e+03  2.397e+03  -1.789  0.0736 .
x1           9.676e-01  4.291e-04 2255.055 <2e-16 ***
x2           9.328e+01  2.116e+00   44.076 <2e-16 ***
x3           1.213e+00  1.492e-02   81.314 <2e-16 ***
x4log       -5.391e+04  5.419e+03  -9.948 <2e-16 ***
x5          -1.381e+03  3.561e+03  -0.388  0.6982
mes         -2.720e+05  1.440e+04  -18.885 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2784000 on 149930 degrees of freedom
Multiple R-squared:  0.987,    Adjusted R-squared:  0.987
F-statistic: 1.628e+06 on 7 and 149930 DF, p-value: < 2.2e-16
```

Tabla 10. Prueba 2

No aumenta significativamente R2

Reglogx5<-lm(luz~x0+x1+x2+x3+x4+x5log+mes,data=Panel1)

summary(Reglogx5)

```
Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + x5log + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-25830369  -218976  -115637   103587 255297556

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.398e+05  7.552e+04   7.148 8.87e-13 ***
x0          -4.200e+03  2.389e+03  -1.758  0.0788 .
x1           9.651e-01  4.350e-04 2218.669 < 2e-16 ***
x2           9.415e+01  2.012e+00  46.782 < 2e-16 ***
x3           1.528e+00  1.794e-02  85.183 < 2e-16 ***
x4          -8.289e+00  2.530e-01 -32.765 < 2e-16 ***
x5log        -7.504e+02  7.188e+03  -0.104  0.9169
mes          -2.683e+05  1.435e+04 -18.690 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2775000 on 149930 degrees of freedom
Multiple R-squared:  0.9871,    Adjusted R-squared:  0.9871
F-statistic: 1.638e+06 on 7 and 149930 DF, p-value: < 2.2e-16
```

Tabla 11. Prueba 3

No aumenta significativamente R2

Opción 2. Eliminando variables

Elimx1<-lm(luz~x0+x2+x3+x4+x5+mes,data=Panel1)

summary(Elimx1)

```
Call:
lm(formula = luz ~ x0 + x2 + x3 + x4 + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-433961478  -1933423  -801723   384587 1411596582

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.671e+05  4.434e+05  -1.730  0.0836 .
x0          -6.775e+03  1.390e+04  -0.487  0.6259
x2           1.120e+03  1.139e+01  98.316 <2e-16 ***
x3           2.621e+01  8.183e-02 320.258 <2e-16 ***
x4          -1.115e+02  1.446e+00 -77.059 <2e-16 ***
x5           6.099e+03  2.065e+04   0.295  0.7677
mes          1.090e+06  8.342e+04  13.069 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16140000 on 149931 degrees of freedom
Multiple R-squared:  0.5635,    Adjusted R-squared:  0.5634
F-statistic: 3.225e+04 on 6 and 149931 DF, p-value: < 2.2e-16
```

Tabla 12. Prueba 4

Disminuye considerablemente R2

Elimx4<-lm(luz~x0+x1+x2+x3+x5+mes,data=Panel1)

summary(Elimx4)


```

Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x5 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-29593056  -217050  -120386   108276 258794875

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.484e+05  7.651e+04   7.169 7.61e-13 ***
x0          -4.305e+03  2.398e+03  -1.795  0.0726 .
x1           9.677e-01  4.291e-04 2255.202 < 2e-16 ***
x2           8.656e+01  2.006e+00  43.145 < 2e-16 ***
x3           1.170e+00  1.428e-02  81.946 < 2e-16 ***
x5          -1.314e+03  3.562e+03  -0.369  0.7123
mes          -2.737e+05  1.440e+04 -18.998 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2785000 on 149931 degrees of freedom
Multiple R-squared:  0.987,    Adjusted R-squared:  0.987
F-statistic: 1.898e+06 on 6 and 149931 DF,  p-value: < 2.2e-16

```

Tabla 13. Prueba 5

No aumenta R2

Elimx4mes<-lm(luz~x0+x1+x2+x3+x5,data=Panel1)

summary(Elimx4mes)

```

Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x5, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-29644132  -152980  -42408   -2092 258711484

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.406e+05  7.352e+04   1.913  0.0558 .
x0          -4.353e+03  2.401e+03  -1.813  0.0698 .
x1           9.674e-01  4.293e-04 2253.636 <2e-16 ***
x2           8.799e+01  2.007e+00  43.837 <2e-16 ***
x3           1.175e+00  1.429e-02  82.228 <2e-16 ***
x5          -1.294e+03  3.567e+03  -0.363  0.7168
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2788000 on 149932 degrees of freedom
Multiple R-squared:  0.987,    Adjusted R-squared:  0.987
F-statistic: 2.272e+06 on 5 and 149932 DF,  p-value: < 2.2e-16

```

Tabla 14. Prueba 6

No aumenta R2

Elimx5<-lm(luz~x0+x1+x2+x3+x4+mes,data=Panel1)

summary(Elimx5)

```

Call:
lm(formula = luz ~ x0 + x1 + x2 + x3 + x4 + mes, data = Panel1)

Residuals:
    Min       1Q   Median       3Q      Max
-25829088 -219008 -115558  103587 255296863

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.391e+05  7.521e+04   7.168 7.67e-13 ***
x0          -4.200e+03  2.389e+03  -1.758  0.0788 .
x1           9.651e-01  4.350e-04 2218.677 < 2e-16 ***
x2           9.415e+01  2.012e+00  46.782 < 2e-16 ***
x3           1.528e+00  1.794e-02  85.183 < 2e-16 ***
x4          -8.289e+00  2.530e-01 -32.765 < 2e-16 ***
mes          -2.683e+05  1.435e+04 -18.690 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2775000 on 149931 degrees of freedom
Multiple R-squared:  0.9871,    Adjusted R-squared:  0.9871
F-statistic: 1.912e+06 on 6 and 149931 DF,  p-value: < 2.2e-16

```

Tabla 15. Prueba 7

No aumenta significativamente R

Con ninguna de las dos alternativas se logra incrementar la R cuadrada, sin embargo se observa que eliminando algunas variables específicas (x1 por ejemplo) dicho valor disminuye significativamente. Sin embargo, considero que la eliminación de variables no es una opción aceptable, mientras que utilizar algunas variables en logaritmo para eliminar el efecto de las unidades de las variables sobre los coeficientes, si lo es, pero no representan estas ningún cambio considerable en el resultado de R cuadrada.

EXAMEN, CODIGO

```

1 #Juan Sebastian Velasquez Diaz
2 #Primer examen Metodos II
3
4
5 rm(list=ls())
6 Base1<-read.csv("examen1.csv")
7
8 install.packages("plm")
9 install.packages("ggplot2")
10 library("plm")
11 library("ggplot2")
12
13
14
15
16 #Punto 1. Incorporar la base de datos (panel) en R, y declararla como panel.
17
18 Panel1 <- pdata.frame(Base1, index = c("clave", "t"))
19 Panel1
20
21 #Punto 2.Realizar una tabla de estadísticas descriptivas de las variables.
22
23 reg1<-lm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1)
24 summary(reg1)
25 head(Base1)
26
27 #Punto 3. Realizar regresion pooled-OLS
28
29 OLS<- lm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1)
30 summary(OLS)

```

```

31
32 #Punto 4. Determinar si la regresión es de efectos fijos o aleatorios
33
34 #efectos fijos
35 ef<-plm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1,index=c("clave", "t"),model="within")
36 summary(ef)
37
38 #efectos aleatorios
39 ea <- plm(luz~x0+x1+x2+x3+x4+x5+mes,data=Panel1, index=c("clave", "t"), model="random")
40 summary(ea)
41
42 #Prueba para elegir efectos fijos o efectos aleatorios
43 phtest(ef, ea)
44
45 #Se decide utilizar efectos fijos porque la prueba HT arroja un valor significativo de P-Value
46
47 #Punto 5. Probar si la base de datos tiene heteroscedasticidad, si la tiene, corregirla
48
49 library(lmtest)
50 bptest(ef)
51
52 #Se comprueba que la muestra si tiene heterocedasticidad
53
54 #Por lo tanto se debe corregir
55
56 vc_ef <- vcovHC(ef, type="HC0")
57 summary(vc_ef)

```

```

58
59 #Punto 6.¿Que otras maneras de presentar las variables se te ocurre
60 #¿Es valida esa especificacion?
61
62 #opcion 1. Haciendo logaritmicas algunas variables
63 Panel1$x0log<-log(Panel1$x0)
64 Panel1$x4log<-log(Panel1$x4)
65 Panel1$x5log<-log(Panel1$x5)
66
67 Reglogx0<-lm(luz~x0log+x1+x2+x3+x4+x5+mes,data=Panel1)
68 summary(Reglogx0)
69 #No aumento R2
70
71 Reglogx4<-lm(luz~x0+x1+x2+x3+x4log+x5+mes,data=Panel1)
72 summary(Reglogx4)
73 #No aumenta signifcativamente R2
74
75 Reglogx5<-lm(luz~x0+x1+x2+x3+x4+x5log+mes,data=Panel1)
76 summary(Reglogx5)
77 #No aumenta significativamente R2
78
79 #Opción 2. Elimando variables
80
81 Elimx1<-lm(luz~x0+x2+x3+x4+x5+mes,data=Panel1)
82 summary(Elimx1)
83 #Disminuye conasiderablemente R2
84

```

```

84
85 Elimx4<-lm(luz~x0+x1+x2+x3+x5+mes,data=Panel1)
86 summary(Elimx4)
87 #No aumenta R2
88
89 Elimx4mes<-lm(luz~x0+x1+x2+x3+x5,data=Panel1)
90 summary(Elimx4mes)
91 #No aumenta R2
92
93 Elimx5<-lm(luz~x0+x1+x2+x3+x4+mes,data=Panel1)
94 summary(Elimx5)
95 #No aumenta significativamente R
96

```