

Modelo de panel espacial para la predicción del VA municipal

Este documento tiene por objetivo la presentación de un modelo de pronóstico para el valor agregado municipal que aprovecha la dimensión geográfica y la dimensión espacial de los datos.

Se inicia cargando las librerías necesarias.

```
library(sp)
library(rgdal)
library(spdep)
library(readxl)
library(tidyverse)
library(splm)
library(panelr)
library("tmap")
```

En los datos, se tienen:

1. El valor agregado del sector educación por municipio, distribuido desde el departamental bajo el supuesto de que la participación del sector educación en el sector terciario es igual para los municipios que para el departamento. Así, desde el valor agregado del sector terciario por municipio se llega al valor agregado de la educación por municipio.
2. El número de establecimientos educativos como un proxy al capital.
3. El número de profesores como proxy a la fuerza laboral.
4. El número de estudiantes como proxy a las transferencias
5. El Shape de Antioquia que contiene los polígonos para cada municipio.

```
shape_antioquia <- readOGR(dsn=file.choose(),layer="MpiosAntioquia")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\josej\Desktop\Gobernación\Tareas\Modelo espacial\MpiosAntioquia.shp", layer: "MpiosAntioquia"
## with 125 features
## It has 13 fields
## Integer64 fields read as strings: OBJECTID
```

```
datos <- read_excel(file.choose())
datos_ent <- datos[,c("Codigo", "alumnos_2015", "docentes_2015",
                     "establecimientos_2015", "alumnos_2016",
                     "docentes_2016", "establecimientos_2016",
                     "PIB_2015", "PIB_2016")]
names(datos_ent)[1] <- "COD_MPIO"
```

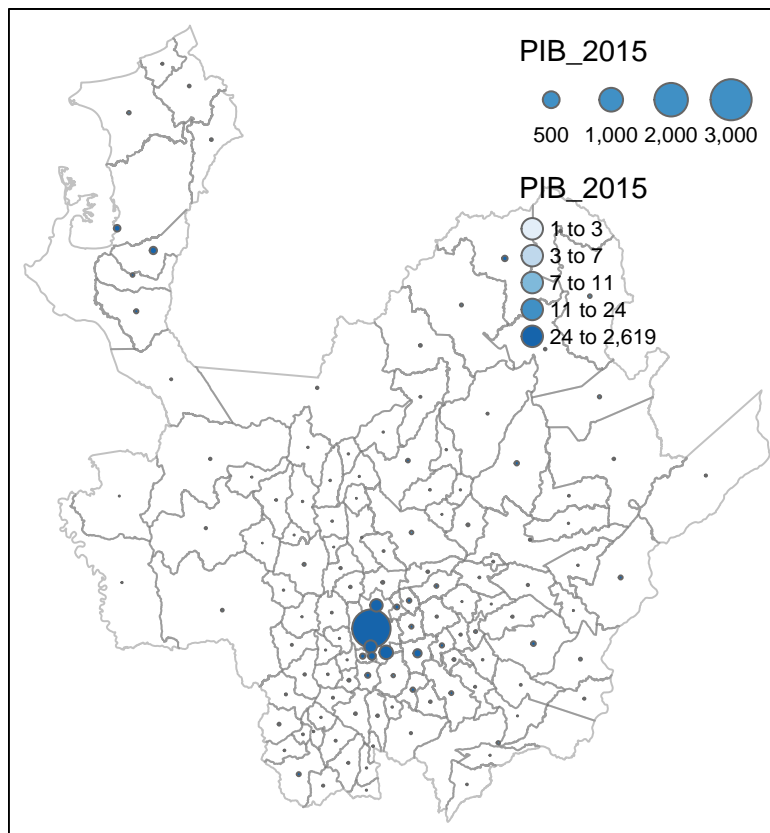
A continuación se visualiza el valor agregado para el 2015 y el 2016 por municipios, aprovechando el shape de Antioquia.

```

coordenadas <- cbind.data.frame (shape_antioquia@data[,2],coordinates(shape_antioquia))
names(coordenadas)[1] <- "Codigo"
pib_mun <- merge(coordenadas,datos[,c(1,14,15)],by="Codigo")
pib_mun <- SpatialPointsDataFrame(pib_mun[,2:3], pib_mun)
proj4string(pib_mun) = proj4string(shape_antioquia)

# Mapa de puntos coloreado
tm_shape(shape_antioquia) + tm_borders(alpha=.4) +
  tm_shape(pib_mun) +
  tm_bubbles(size = "PIB_2015", col = "PIB_2015", palette = "Blues", style = "quantile")

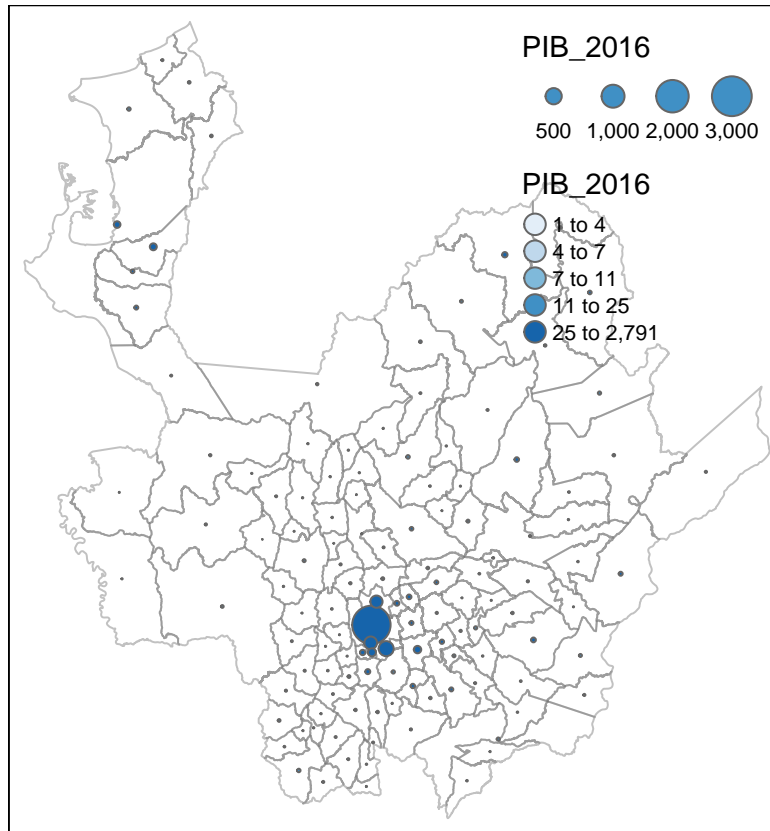
```



```

tm_shape(shape_antioquia) + tm_borders(alpha=.4) +
  tm_shape(pib_mun) +
  tm_bubbles(size = "PIB_2016", col = "PIB_2016", palette = "Blues", style = "quantile")

```



Para la estimación del modelo, los datos se convierten en formato largo, es decir, donde las mediciones de las variables en el tiempo se incorporan en una misma columna; esto es contrario al formato ancho donde para cada año existe una columna que contiene la medición de la variable para cada municipio.

```
datos_ent <- long_panel(datos_ent,prefix="_20",begin=15,end=16,label_location="end",wave="Año")
```

Para la estimación de todo modelo espacial se debe introducir una matriz de pesos espaciales, que define la relación de vecindad de los municipios.

```
nb_antioquia <- poly2nb(shape_antioquia)
pesos_antioquia <- nb2listw(nb_antioquia)
```

El modelo espacial, siguiendo a Lampis (2016), intenta aproximarse a la estimación del valor agregado desde la perspectiva de una función de producción; es decir, el valor agregado es una función del capital y del trabajo usados en el sector particular; no obstante, ya que también se incluye educación de no mercado, se debe incluir una variable que capture el efecto de las transferencias del estado sobre el valor de producción del sector. Este modelo espacial puede ser de dos tipos: un modelo de rezagos espaciales, donde el valor agregado de un municipio tiene asociaciones positivas o negativas con el de sus vecinos, y un modelo de errores espaciales, donde es en el término de error donde se presentan dichas asociaciones. En este caso se estiman ambos modelos y se verifica la significancia de los parámetros; este es conocido como un modelo Sarar.

Respecto a la dimensión espacial, ya que se siguen múltiples variables para múltiples unidades muestrales en el tiempo, se tiene un modelo de panel de datos. Los modelos de panel de datos son principalmente de dos tipos: los modelos de efectos fijos y los modelos de efectos aleatorios. La decisión sobre cuál modelo utilizar depende enteramente de la covarianza entre el término de error y el término de efecto individual no

observado de los datos, esto se verifica mediante la prueba de Hausman. Según Arbia (2014), la aceptación de la hipótesis nula implica que el modelo a utilizar debe ser el de efectos aleatorios, y viceversa.

Al realizar la prueba, se obtiene un rechazo a la hipótesis nula, de manera que el modelo correcto es el modelo de efectos fijos.

```
sphtest(PIB~alumnos+establecimientos+docentes,
        data=datos_ent, listw=pesos_antioquia,
        spatial.model="sarar",
        method="ML", index="id",)

##
## Hausman test for spatial models
##
## data: x
## chisq = 545.65, df = 3, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

Arbia (2014) presenta los métodos de estimación de los modelos de panel espaciales. El método de estimación correcto para un modelo Sarar de efectos fijos es la máxima verosimilitud, la cual se puede implementar mediante la función spml.

En la parte inferior puede observarse que el término “lambda” no es estadísticamente significativo lo cual indica que no hay un efecto de correlación entre el valor agregado de un municipio y sus vecinos, no obstante este efecto si existe para los errores.

```
modelo <- spml(PIB~alumnos+establecimientos+docentes,
               data=datos_ent, index=c("id"),
               spatial.error="b", listw=pesos_antioquia,
               model="within", effect="individual", lag=TRUE)

summary(modelo)

## Spatial panel fixed effects sarar model
##
##
## Call:
## spml(formula = PIB ~ alumnos + establecimientos + docentes, data = datos_ent,
##       index = c("id"), listw = pesos_antioquia, model = "within",
##       effect = "individual", lag = TRUE, spatial.error = "b")
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -1.5347e+01 -4.5255e-01  3.0670e-15  4.5255e-01  1.5347e+01
##
## Spatial error parameter:
##      Estimate Std. Error t-value Pr(>|t|)
## rho 0.223026   0.094795  2.3527  0.01864 *
##
## Spatial autoregressive coefficient:
##      Estimate Std. Error t-value Pr(>|t|)
## lambda 0.011460   0.034636  0.3309  0.7407
##
## Coefficients:
```

```
##               Estimate Std. Error t-value Pr(>|t|)
## alumnos        0.00218122  0.00013428  16.244 < 2.2e-16 ***
## establecimientos 0.95962375  0.06489577  14.787 < 2.2e-16 ***
## docentes       -0.06028986  0.00275873 -21.854 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

El modelo se reestima sin el componente SAR:

```
modelo <- spml(PIB~alumnos+establecimientos+docentes,
  data=datos_ent,index=c("id"),
  spatial.error="b",listw=pesos_antioquia,
  model="within",effect="individual")

summary(modelo)
```

```
## Spatial panel fixed effects error model
##
##
## Call:
## spml(formula = PIB ~ alumnos + establecimientos + docentes, data = datos_ent,
##       index = c("id"), listw = pesos_antioquia, model = "within",
##       effect = "individual", spatial.error = "b")
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -1.5342e+01 -4.5122e-01 -3.5527e-15  4.5122e-01  1.5342e+01
##
## Spatial error parameter:
##      Estimate Std. Error t-value Pr(>|t|)
## rho 0.237344  0.088704  2.6757 0.007458 **
##
## Coefficients:
##               Estimate Std. Error t-value Pr(>|t|)
## alumnos        0.0021822  0.0001342  16.260 < 2.2e-16 ***
## establecimientos 0.9579346  0.0646258  14.823 < 2.2e-16 ***
## docentes       -0.0602438  0.0027566 -21.854 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Los efectos individuales del modelo que recogen los efectos de todas las variables faltantes para explicar el valor agregado municipal se pueden recuperar mediante la función `effects`.

```
efectos <- effects(modelo)

efectos[[2]]
```

```
##      Estimate Std. Error      t-value      Pr(>|t|)
## 1  3083.09054098 133.864102  23.03149604 2.254708e-117
## 2   -20.31090962   1.821729 -11.14925086 7.221175e-29
## 3    84.82678571   7.237662  11.72019225 1.004438e-31
## 4   -0.90771253   2.718331  -0.33392272 7.384379e-01
```

## 5	-17.90404178	1.921272	-9.31884985	1.176082e-20
## 6	309.04322327	4.965379	62.23960192	0.000000e+00
## 7	-13.46006637	1.791913	-7.51156260	5.842572e-14
## 8	124.65480773	4.370208	28.52376860	5.943450e-179
## 9	-28.11533816	1.882656	-14.93386755	1.984163e-50
## 10	49.68159445	2.690390	18.46631920	3.854894e-76
## 11	-35.99040661	1.796984	-20.02823309	3.125521e-89
## 12	-15.80578941	3.324049	-4.75498153	1.984646e-06
## 13	-48.01313919	1.944570	-24.69087418	1.340322e-134
## 14	-42.08884080	1.768287	-23.80203852	3.181325e-125
## 15	-40.22401055	1.800013	-22.34650935	1.305604e-110
## 16	-39.41972836	1.819863	-21.66082613	4.805391e-104
## 17	-35.71363290	1.704436	-20.95334040	1.749208e-97
## 18	-34.95126116	1.706605	-20.47999542	3.246890e-93
## 19	-24.83575936	1.915217	-12.96759376	1.868029e-38
## 20	-26.89492185	1.760466	-15.27715850	1.085726e-52
## 21	-33.46862699	1.737963	-19.25738946	1.223956e-82
## 22	-35.00356960	1.724143	-20.30201157	1.234275e-91
## 23	-5.54672904	1.760077	-3.15141192	1.624832e-03
## 24	-35.37208071	1.719104	-20.57588232	4.514402e-94
## 25	-31.41809871	1.703442	-18.44388777	5.838806e-76
## 26	-33.43717098	1.774018	-18.84826625	3.036314e-79
## 27	-30.75050006	1.740754	-17.66504883	7.794534e-70
## 28	-33.92679328	1.719255	-19.73342398	1.113658e-86
## 29	-27.90858316	1.807521	-15.44025803	8.774852e-54
## 30	-30.92422639	1.721154	-17.96714669	3.523787e-72
## 31	-35.23842871	1.701528	-20.70987100	2.822112e-95
## 32	-33.24972545	1.747330	-19.02887826	9.833517e-81
## 33	-34.57785599	1.708509	-20.23861806	4.475430e-91
## 34	-38.19736769	1.721845	-22.18397704	4.904598e-109
## 35	-38.05481972	1.724088	-22.07243287	5.817559e-108
## 36	-35.65429991	1.706255	-20.89622765	5.794747e-97
## 37	-34.38812071	1.696923	-20.26498564	2.620370e-91
## 38	-32.72377895	1.725741	-18.96215575	3.504800e-80
## 39	-33.88691026	1.705475	-19.86948492	7.476608e-88
## 40	-21.61821531	1.703853	-12.68783873	6.906541e-37
## 41	-35.97843968	1.700837	-21.15336933	2.569199e-99
## 42	-39.91823341	1.906925	-20.93329899	2.664054e-97
## 43	-43.36639487	1.804682	-24.02993514	1.353372e-127
## 44	-34.96999511	1.696494	-20.61309624	2.093993e-94
## 45	-30.15048834	1.744511	-17.28306166	6.310737e-67
## 46	-33.24892094	2.160763	-15.38758528	1.982954e-53
## 47	-34.96066695	1.701958	-20.54144506	9.179055e-94
## 48	-37.52304236	1.725274	-21.74902805	7.055876e-105
## 49	-29.30228535	1.872768	-15.64651252	3.509312e-55
## 50	-35.61525327	1.697151	-20.98531886	8.932200e-98
## 51	-36.65017927	1.706082	-21.48207526	2.290420e-102
## 52	-34.27374273	1.696616	-20.20123441	9.548051e-91
## 53	-35.62274327	1.703373	-20.91306303	4.072318e-97
## 54	-35.28381238	1.701384	-20.73829443	1.563725e-95
## 55	-34.47237219	1.731697	-19.90669944	3.560110e-88
## 56	-38.72634484	1.768072	-21.90315141	2.424204e-106
## 57	-33.22798123	1.716317	-19.36004860	1.677272e-83
## 58	-36.21682177	1.776923	-20.38175899	2.427724e-92

## 59	-37.76838413	1.709971	-22.08715391	4.200397e-108
## 60	-35.36363908	1.703493	-20.75948640	1.006364e-95
## 61	-34.32977049	1.713188	-20.03853111	2.541579e-89
## 62	-35.60861818	1.700325	-20.94224382	2.208115e-97
## 63	-34.63844402	1.707633	-20.28447219	1.763455e-91
## 64	-34.50555028	1.709692	-20.18232492	1.400039e-90
## 65	-29.93621503	1.719660	-17.40821533	7.147761e-68
## 66	-19.02120900	1.947383	-9.76757406	1.551219e-22
## 67	-30.43747150	1.719429	-17.70207739	4.041213e-70
## 68	-35.59120286	1.702307	-20.90763257	4.563198e-97
## 69	-31.46860455	1.744275	-18.04107789	9.271770e-73
## 70	-34.44188221	1.696170	-20.30567743	1.145536e-91
## 71	-36.80094461	1.713092	-21.48217625	2.285446e-102
## 72	-37.73514656	1.788230	-21.10194951	7.632364e-99
## 73	-34.41800190	1.696418	-20.28862921	1.620507e-91
## 74	46.95041209	4.354482	10.78208877	4.182781e-27
## 75	-29.81020415	1.719713	-17.33440272	2.587410e-67
## 76	-31.58215842	1.752650	-18.01965704	1.365876e-72
## 77	-33.33827118	1.753528	-19.01211370	1.353858e-80
## 78	-34.09387263	1.710076	-19.93705183	1.941768e-88
## 79	-17.52793273	1.805528	-9.70792608	2.789551e-22
## 80	-19.72026657	1.698744	-11.60873561	3.720767e-31
## 81	-24.56168579	1.841065	-13.34101798	1.336223e-40
## 82	-33.28138610	1.728043	-19.25957999	1.173263e-82
## 83	-20.20514090	1.891224	-10.68362979	1.214289e-26
## 84	-34.14622354	1.713368	-19.92929425	2.267367e-88
## 85	30.52120758	5.016265	6.08444940	1.168923e-09
## 86	0.08944518	1.727430	0.05177934	9.587045e-01
## 87	-35.97121970	1.712718	-21.00242190	6.232595e-98
## 88	-38.50690823	1.761361	-21.86201669	5.974555e-106
## 89	-28.93121193	1.713086	-16.88835763	5.480753e-64
## 90	-30.71892681	1.725626	-17.80161769	6.865670e-71
## 91	-21.93983322	1.844058	-11.89758565	1.218207e-32
## 92	-25.48177043	1.747349	-14.58310684	3.597642e-48
## 93	-0.50258422	2.408965	-0.20863075	8.347365e-01
## 94	-35.85270465	1.702039	-21.06455649	1.681935e-98
## 95	-33.92032383	1.704101	-19.90511647	3.674368e-88
## 96	-35.35542821	1.717078	-20.59045767	3.341972e-94
## 97	-34.15641272	1.700586	-20.08508927	9.964413e-90
## 98	-29.00065898	1.752203	-16.55097120	1.575406e-61
## 99	-35.22097595	1.730220	-20.35635191	4.078437e-92
## 100	-30.42424375	1.772813	-17.16156351	5.150812e-66
## 101	-33.66287777	1.698119	-19.82362828	1.861885e-87
## 102	-32.05676990	1.713445	-18.70895817	4.184757e-78
## 103	-29.11126515	1.714693	-16.97753940	1.204304e-64
## 104	-34.40293505	1.702280	-20.20991252	8.008980e-91
## 105	-34.39952501	1.698331	-20.25489564	3.216308e-91
## 106	-34.59717358	1.699793	-20.35375603	4.300279e-92
## 107	-33.22909705	1.722376	-19.29258903	6.198986e-83
## 108	-29.61402142	1.735990	-17.05886420	3.003381e-65
## 109	-31.95746630	1.724839	-18.52779736	1.232290e-76
## 110	-36.07669908	1.706965	-21.13499765	3.792075e-99
## 111	-34.58424028	1.714196	-20.17519772	1.617157e-90
## 112	-28.22118603	1.810480	-15.58768393	8.827646e-55

```
## 113 -35.36571635 1.706321 -20.72630257 2.006246e-95
## 114 -34.47120652 1.732214 -19.90008579 4.062328e-88
## 115 -39.11740755 3.959107 -9.88036166 5.064978e-23
## 116 -37.35295075 1.838923 -20.31240058 9.990033e-92
## 117 -43.73362088 1.909622 -22.90171496 4.466906e-116
## 118 -37.55426240 1.979657 -18.97008541 3.014165e-80
## 119 -38.78047090 1.704075 -22.75749497 1.209618e-114
## 120 -37.51553429 1.732659 -21.65199492 5.820562e-104
## 121 -44.49892079 2.065269 -21.54631051 5.733796e-103
## 122 -36.04210259 1.790151 -20.13355128 3.751230e-90
## 123 -45.60823112 1.835905 -24.84237496 3.126205e-136
## 124 29.37354932 5.410579 5.42891032 5.669916e-08
## 125 -40.47327827 1.728490 -23.41539193 2.978757e-121
```

La función splm no contiene un método de predicción por lo que esta se debe hacer de manera manual.

```
datos_pron <- datos[,c("Codigo", "alumnos_2017", "docentes_2017", "establecimientos_2017")]

efectos[[1]][1]+efectos[[2]][,1]+
  datos_pron$alumnos_2017*summary(modelo)$coeff[2]+
  datos_pron$docentes_2017*summary(modelo)$coeff[4]+
  datos_pron$establecimientos_2017*summary(modelo)$coeff[3]
```

```
##      1      2      3      4      5      6
## 2332.8586717 27.6010473 296.2631522 56.1683908 46.1681329 329.1670497
##      7      8      9     10     11     12
## 41.7783188 261.9552442 53.7962676 115.3478336 15.6563931 48.2153317
##     13     14     15     16     17     18
## 17.4664682 10.3805153 11.7147452 14.1620805 2.2880752 3.1154824
##     19     20     21     22     23     24
## 21.7166449 11.0439675 5.7458638 8.5189314 34.6446065 6.0901369
##     25     26     27     28     29     30
## 5.3820524 10.3357895 8.7493123 5.4982747 18.3294705 5.6276280
##     31     32     33     34     35     36
## 1.3794311 9.5479049 5.6699711 1.5220230 -1.0416369 2.6507694
##     37     38     39     40     41     42
## 2.3531863 12.2965915 6.7782245 16.9607171 2.6870778 3.3372478
##     43     44     45     46     47     48
## -4.9070153 1.1656442 14.6215854 19.7097702 2.8638966 2.9581263
##     49     50     51     52     53     54
## 24.0196160 1.1773338 2.9078072 3.3383018 3.5668114 3.2898208
##     55     56     57     58     59     60
## 5.2078139 6.0210250 5.3081222 8.3137213 1.5700589 2.7730094
##     61     62     63     64     65     66
## 2.5863838 1.0424393 2.2268175 4.2969348 9.2752763 16.3504219
##     67     68     69     70     71     72
## 7.2541736 3.0602091 6.6441172 2.5608972 0.7957715 2.7942974
##     73     74     75     76     77     78
## 2.4656146 17.9345079 10.1541867 22.9329884 17.8666747 3.4367556
##     79     80     81     82     83     84
## 30.2780189 18.0864496 37.0497550 10.7543186 37.7984616 4.3143607
##     85     86     87     88     89     90
## 115.2545706 42.2119174 1.5081685 3.0993207 11.5176113 9.0811673
```


##	91	92	93	94	95	96
##	20.1128140	16.3313039	25.3836942	1.7013453	3.4587638	6.4596787
##	97	98	99	100	101	102
##	3.2695106	13.5988946	8.4158109	10.9220647	4.3522359	7.4637345
##	103	104	105	106	107	108
##	7.7357616	5.7212177	3.4557441	4.5318248	6.5058866	11.2465725
##	109	110	111	112	113	114
##	6.3933114	1.7964175	3.8500002	13.2697019	1.0035291	4.3758145
##	115	116	117	118	119	120
##	29.8097988	14.2518347	23.6662706	30.2161628	0.8756016	5.8899215
##	121	122	123	124	125	
##	19.6220158	9.1553799	12.8097789	79.9948672	0.8303719	

El pronóstico negativo en algunos municipios puede tener su explicación en errores en la medición de las variables, o en el tamaño de la dimensión temporal del panel: algunas variables sufrieron reducciones mientras que el valor agregado aumento, de manera que la relación entre las variables y el valor agregado fue capturada con un signo negativo. Es posible que al aumentar la dimensión temporal, el efecto negativo se corrija.

Referencias

- Lampis, F. (2016). Forecasting the sectoral GVA of a small Spanish region. *Economics and Business Letters*, 5(2), 38-44.
- Arbia, G. (2014). *A primer for spatial econometrics with applications in R*. Palgrave Macmillan.