



UNIVERSIDAD NACIONAL DE COLOMBIA

Modelado de datos en áreas

Estadística espacial

Daniela Arbeláez Montoya
Jefferson Gamboa Betancur
Jean Paul Piedrahita García

Universidad Nacional de Colombia
Ciencias, Escuela de Estadística
Medellín, Colombia
2021

Índice

1. Introducción	2
2. Vecinos espaciales y peso espacial	7
2.1. Objetos vecinos	7
2.2. Objetos de ponderaciones espaciales	7
2.3. Manejo de objetos de ponderaciones espaciales	7
2.4. Uso de pesos para simular la autocorrelación espacial	7
3. Prueba de autocorrelación espacial	7
3.1. Pruebas globales	7
3.2. Pruebas locales	7
4. Ajuste de modelos de datos de área	7
4.1. Enfoques de estadística espacial	7
4.1.1. Modelos autorregresivos simultáneos	7
4.1.2. Modelos autorregresivos condicionales	7
4.1.3. Ajuste de modelos de regresión espacial	7

1. Introducción

A lo largo del desarrollo de este documento, se mostrará la construcción de vecinos y los pesos que se pueden aplicar a los vecindarios. Una vez que este importante y a menudo exigente prerequisite esté en su lugar, se procede a buscar formas de medir la autocorrelación espacial.

Si bien las pruebas se basan en modelos de procesos espaciales, primero se examinan las pruebas y solo posteriormente se pasa al modelado. También es interesante mostrar cómo se puede introducir la autocorrelación espacial en datos independientes, de modo que se puedan realizar simulaciones.

El conjunto de datos con el que se trabajará en esta ocasión, contiene 281 distritos censales para ocho condados centrales del estado de Nueva York complementado con límites de tramo. El área tiene una extensión de unos 160 km de norte a sur y de 120 km de este a oeste.

```
> library(rgdal)
> library(sf)
> library(spdep)
>
> NY8 <- readOGR("Base de datos", "NY8_utm18")

## OGR data source with driver: ESRI Shapefile
## Source: "G:\Mi unidad\Universidad Nacional de Colombia\13. Treceavo semestre\Estadística\NY8_utm18.shp"
## with 281 features
## It has 17 fields

> TCE <- readOGR("Base de datos", "TCE")

## OGR data source with driver: ESRI Shapefile
## Source: "G:\Mi unidad\Universidad Nacional de Colombia\13. Treceavo semestre\Estadística\TCE.shp"
## with 11 features
## It has 5 fields

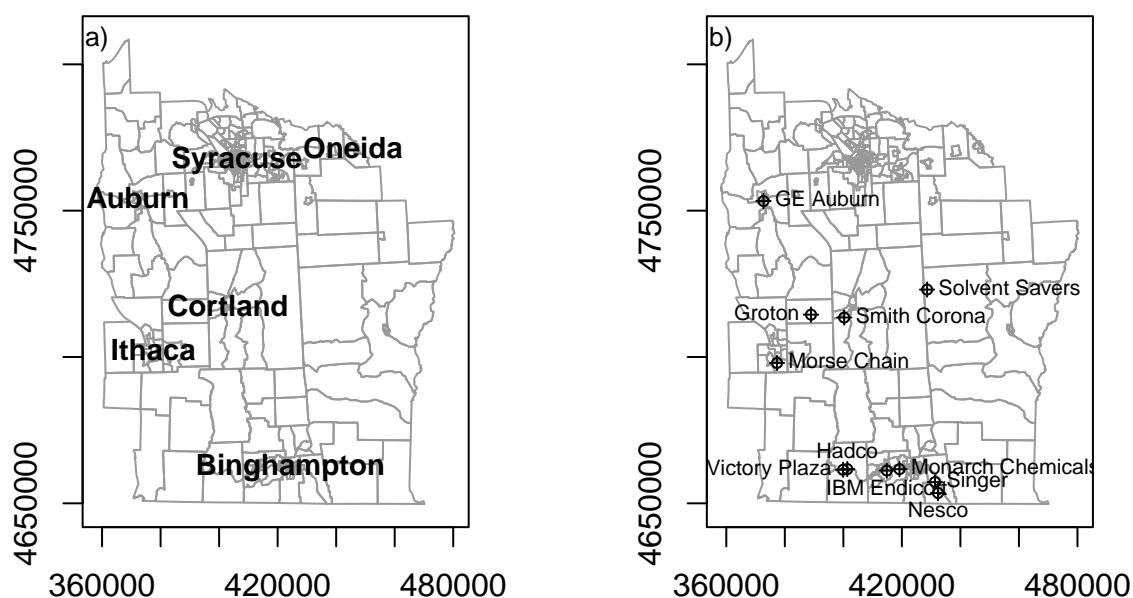
> cities <- readOGR("Base de datos", "NY8cities")

## OGR data source with driver: ESRI Shapefile
## Source: "G:\Mi unidad\Universidad Nacional de Colombia\13. Treceavo semestre\Estadística\NY8cities.shp"
## with 6 features
## It has 1 fields
```

```

> par(mfrow=c(1,2))
> plot(NY8, border="grey60", axes=TRUE)
> text(coordinates(cities), labels=as.character(cities$names), font=2, cex=0.9)
> text(bbox(NY8)[1,1], bbox(NY8)[2,2], labels="a)", cex=0.8)
> plot(NY8, border="grey60", axes=TRUE)
> points(TCE, pch=1, cex=0.7)
> points(TCE, pch=3, cex=0.7)
> text(coordinates(TCE), labels=as.character(TCE$name), cex=0.7,
+ font=1, pos=c(4,1,4,1,4,4,4,2,3,4,2), offset=0.3)
> text(bbox(NY8)[1,1], bbox(NY8)[2,2], labels="b)", cex=0.8)

```

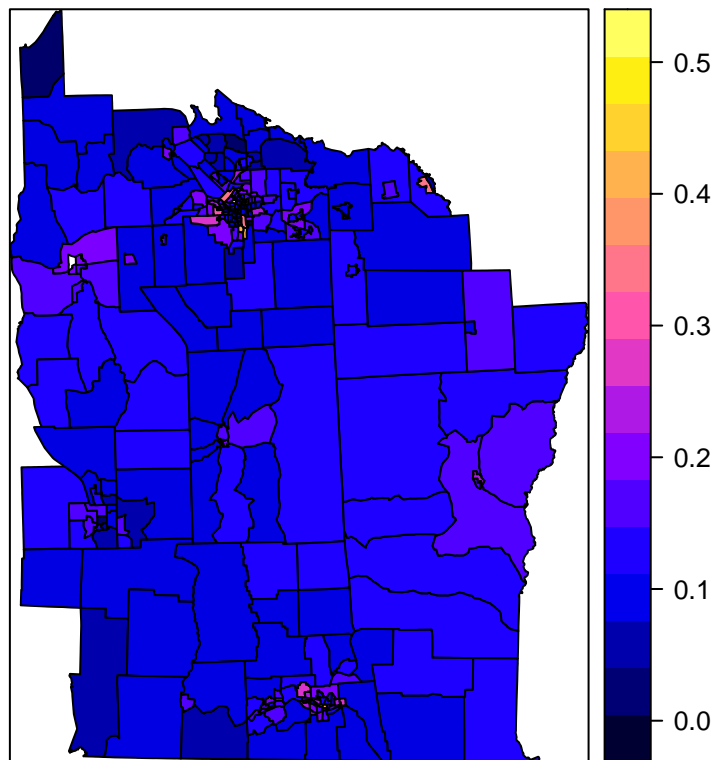


La figura a) muestra las principales ciudades en el área de estudio y b) la ubicación de 11 sitios de desechos peligrosos.

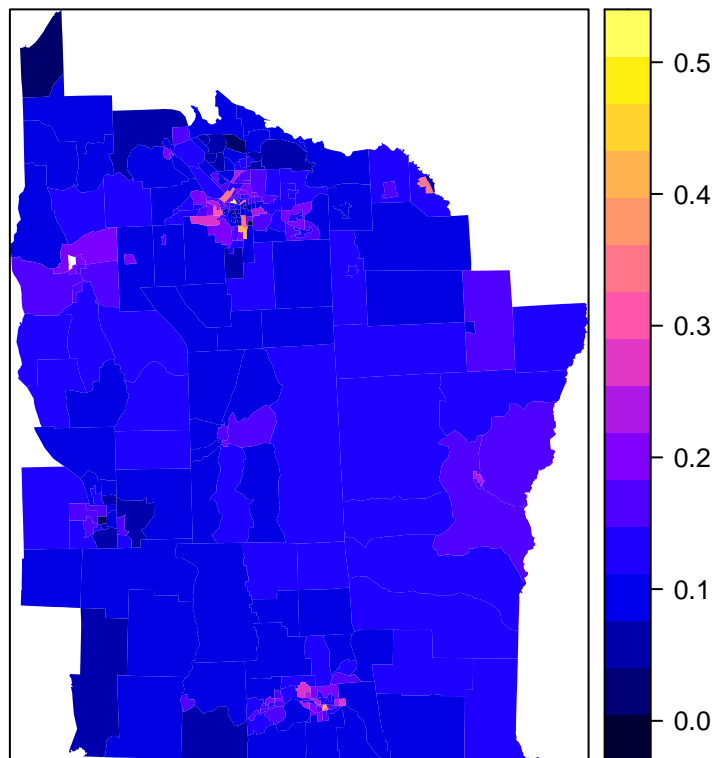
```

> spplot(NY8, c("PCTAGE65P"))

```



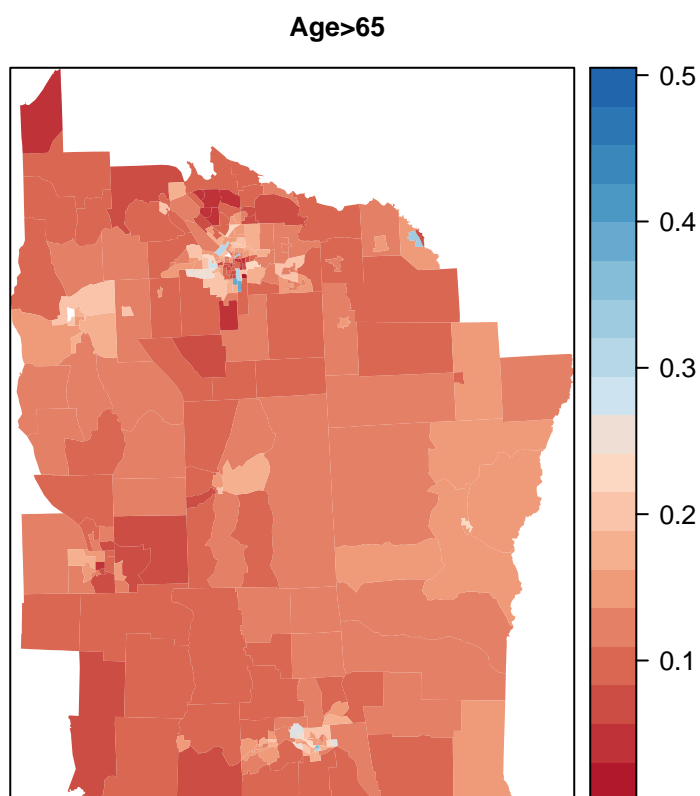
```
> spplot(NY8, c("PCTAGE65P"), col="transparent")
```



```

> library("RColorBrewer")
> #color palette creator function
> rds <- colorRampPalette(brewer.pal(8, "RdBu"))
> #get a range for the values
> tr_at <- seq(min(NY8$PCTAGE65P), max(NY8$PCTAGE65P), length.out=20)
> #create a color interpolating function taking the required
> #number of shades as argument
> tr_rds <- rds(20)
> #parameters
> # at - at which values colors change
> # col.regions - specify fill colors
> tr_pl <- spplot(NY8, c("PCTAGE65P"), at=tr_at, col="transparent",
+               col.regions=tr_rds, main=list(label="Age>65", cex=0.8))
> plot(tr_pl)

```



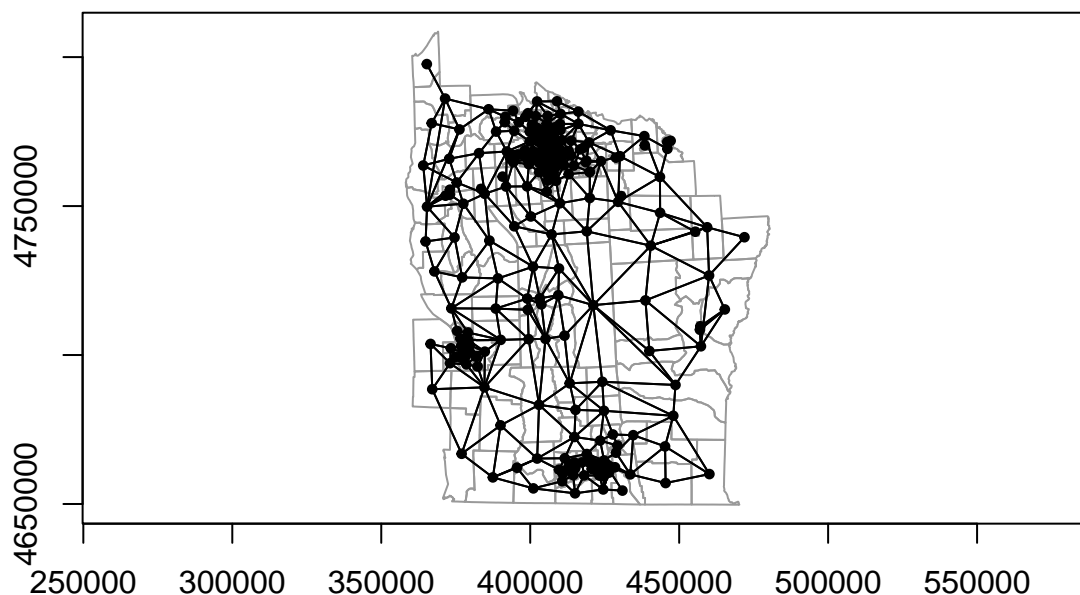
```

> # reads a GAL lattice file into a neighbors list
> NY_nb <- read.gal("Base de datos/NY_nb.gal", region.id = row.names(NY8))
> summary(NY_nb)

```

```
## Neighbour list object:
## Number of regions: 281
## Number of nonzero links: 1522
## Percentage nonzero weights: 1.927534
## Average number of links: 5.41637
## Link number distribution:
##
## 1  2  3  4  5  6  7  8  9 10 11
## 6 11 28 45 59 49 45 23 10  3  2
## 6 least connected regions:
## 55 97 100 101 244 245 with 1 link
## 2 most connected regions:
## 34 82 with 11 links
```

```
> par(mfrow=c(1,1))
> plot(NY8, border="grey60", axes=TRUE)
> plot(NY_nb, coordinates(NY8), pch=19, cex=0.6, add=TRUE)
```



2. Vecinos espaciales y peso espacial

2.1. Objetos vecinos

2.2. Objetos de ponderaciones espaciales

2.3. Manejo de objetos de ponderaciones espaciales

2.4. Uso de pesos para simular la autocorrelación espacial

3. Prueba de autocorrelación espacial

3.1. Pruebas globales

3.2. Pruebas locales

4. Ajuste de modelos de datos de área

4.1. Enfoques de estadística espacial

4.1.1. Modelos autorregresivos simultáneos

4.1.2. Modelos autorregresivos condicionales

4.1.3. Ajuste de modelos de regresión espacial