Spatial Analysis: Columbus Crime

Mary Isangediok

2022-09-20

Spatial Analysis

```
rm(list=ls())
library(spdep)
```

```
## Loading required package: sp
```

```
## Loading required package: spData
```

```
## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source')`
```

```
## Loading required package: sf
```

```
## Linking to GEOS 3.9.1, GDAL 3.4.3, PROJ 7.2.1; sf_use_s2() is TRUE
```

```
library(spatialreg)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'spatialreg'
```

```
## The following objects are masked from 'package:spdep':
##
##     get.ClusterOption, get.coresOption, get.mcOption,
##     get.VerboseOption, get.ZeroPolicyOption, set.ClusterOption,
##     set.coresOption, set.mcOption, set.VerboseOption,
##     set.ZeroPolicyOption
```

```
require(maptools)
```

```
## Loading required package: maptools
```

```
## Checking rgeos availability: TRUE
## Please note that 'maptools' will be retired by the end of 2023,
## plan transition at your earliest convenience;
## some functionality will be moved to 'sp'.
```

```
require(rgdal)
```

```
## Loading required package: rgdal
```

```
## Please note that rgdal will be retired by the end of 2023,
## plan transition to sf/stars/terra functions using GDAL and PROJ
## at your earliest convenience.
##
## rgdal: version: 1.5-32, (SVN revision 1176)
## Geospatial Data Abstraction Library extensions to R successfully loaded
## Loaded GDAL runtime: GDAL 3.4.3, released 2022/04/22
## Path to GDAL shared files: C:/Users/maryp/AppData/Local/R/win-library/4.2/rgdal/gdal
## GDAL binary built with GEOS: TRUE
## Loaded PROJ runtime: Rel. 7.2.1, January 1st, 2021, [PJ_VERSION: 721]
## Path to PROJ shared files: C:/Users/maryp/AppData/Local/R/win-library/4.2/rgdal/proj
## PROJ CDN enabled: FALSE
## Linking to sp version:1.5-0
## To mute warnings of possible GDAL/OSR exportToProj4() degradation,
## use options("rgdal_show_exportToProj4_warnings"="none") before loading sp or rgdal.
```

Load data

Set work directory and plot data

```
setwd("C:\\Users\\maryp\\Downloads\\Spatial_Analysis_Columbus_Crime")

unzip("columbus.zip")
CC = readOGR(dsn = ".", layer = "columbus")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\maryp\Downloads\Spatial_Analysis_Columbus_Crime", layer: "columbus"
## with 49 features
## It has 20 fields
## Integer64 fields read as strings:  COLUMBUS_ COLUMBUS_I POLYID
```

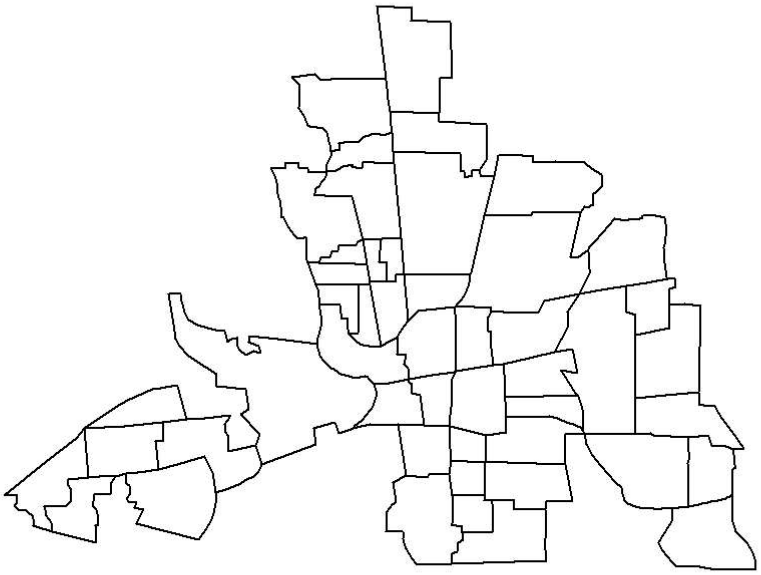
```
## check data
slotNames(CC)
```

```
## [1] "data"        "polygons"    "plotOrder"   "bbox"        "proj4string"
```

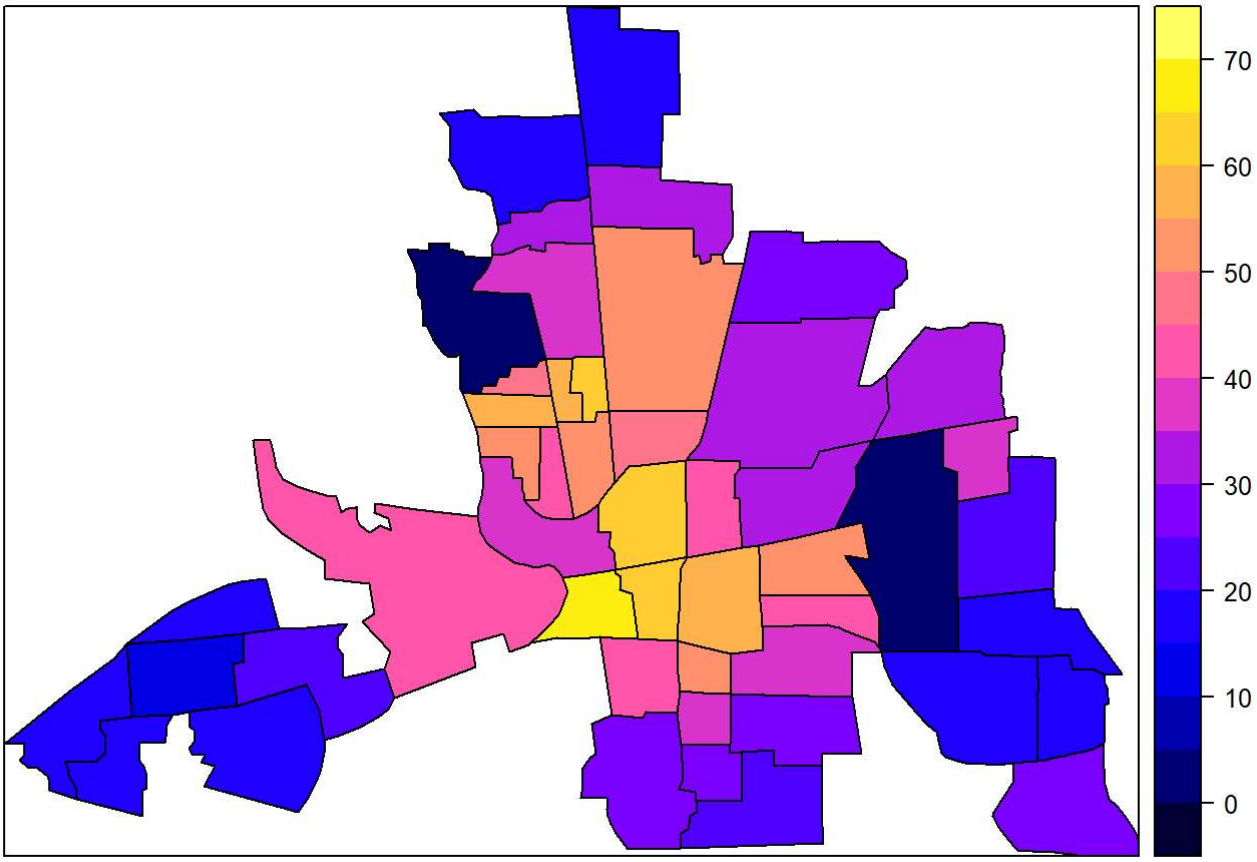
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Plot spatial data

```
df = CC@data
## SSPLOT
plot(CC)
```



```
spplot(CC, "CRIME", pretty = TRUE)
```



Plot Columbus crime data

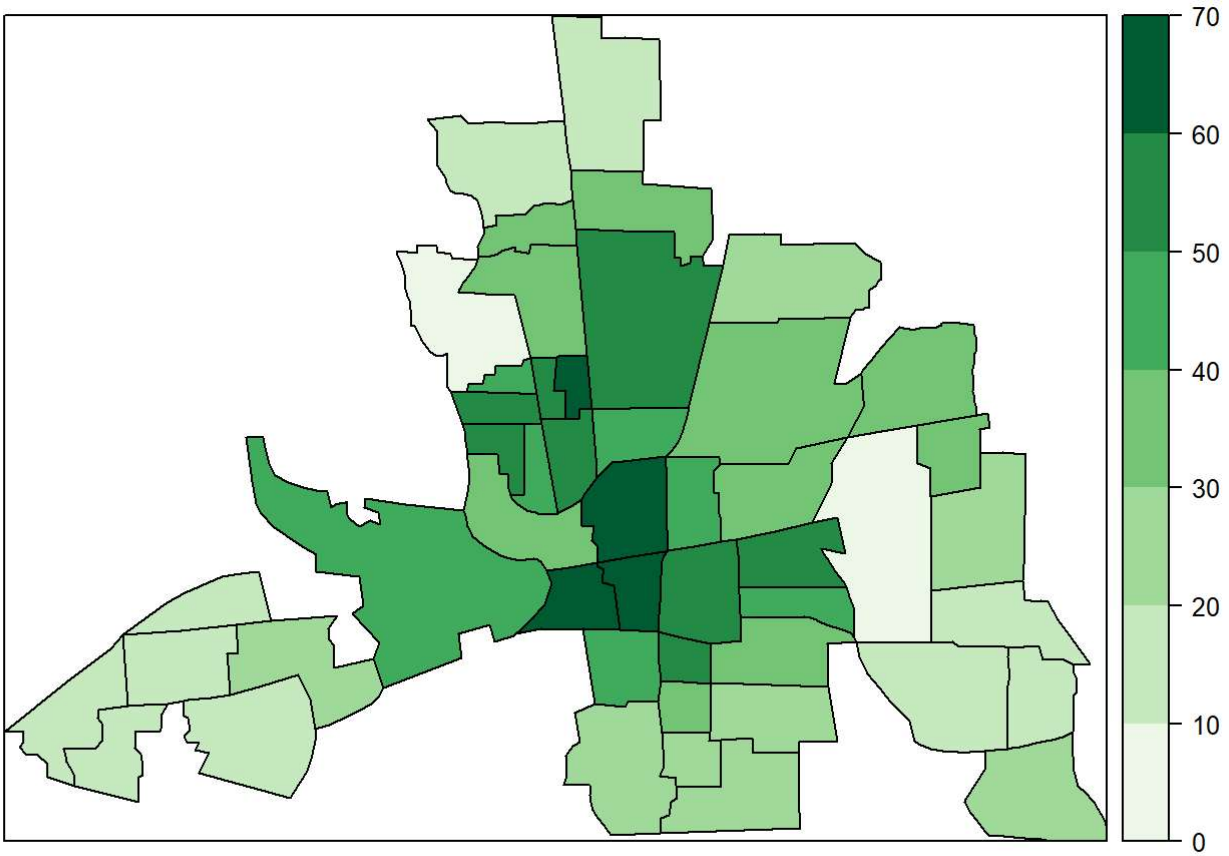
```
require(RColorBrewer)
```

```
## Loading required package: RColorBrewer
```

```
range(CC$CRIME)
```

```
## [1]  0.178269 68.892044
```

```
rng = seq(0, 70, 10)
cls = brewer.pal(7, "Greens")
spplot(CC, "CRIME", col.regions = cls, at = rng, sub = "Crime Data in Ohio")
```



Crime Data in Ohio

Plot neighbor coordinates for data

```
data(columbus)
help(columbus)
```

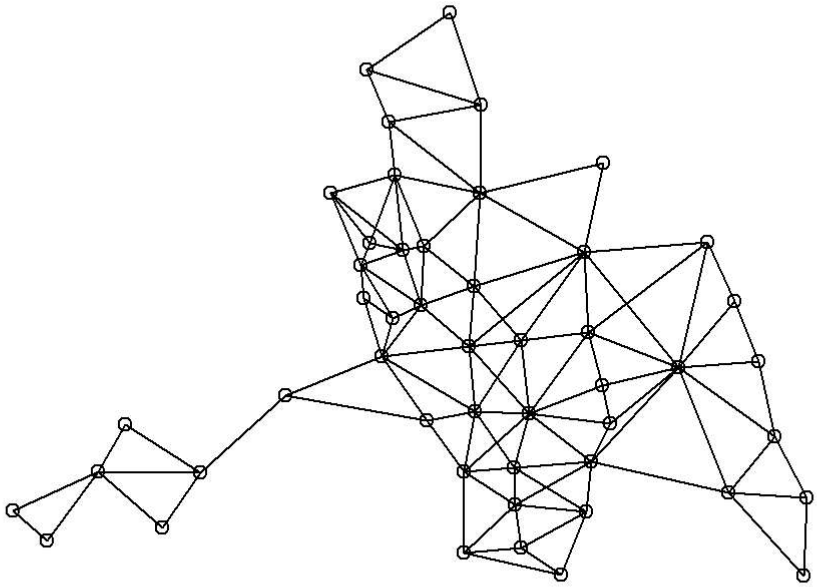
```
## starting httpd help server ... done
```

```
mydata <- columbus
attach(mydata)
Y <- cbind(CRIME)
X <- cbind(INC, HOVAL)
xy <- cbind(mydata$X, mydata$Y)
neighbors <- col.gal.nb
coords <- coords

# Neighbors summary
summary(neighbors)
```

```
## Neighbour list object:
## Number of regions: 49
## Number of nonzero links: 230
## Percentage nonzero weights: 9.579342
## Average number of links: 4.693878
## Link number distribution:
##
##  2  3  4  5  6  7  8  9 10
##  7  7 13  4  9  6  1  1  1
## 7 least connected regions:
## 1005 1008 1045 1047 1049 1048 1015 with 2 links
## 1 most connected region:
## 1017 with 10 links
```

```
plot(neighbors, coords)
```



##

Descriptive statistics

```
summary(Y)
```

```
##      CRIME
##  Min.   : 0.1783
##  1st Qu.:20.0485
##  Median :34.0008
##  Mean   :35.1288
##  3rd Qu.:48.5855
##  Max.   :68.8920
```

```
summary(X)
```

```
##       INC              HOVAL
##  Min.   : 4.477   Min.   :17.90
##  1st Qu.: 9.963   1st Qu.:25.70
##  Median :13.380   Median :33.50
##  Mean   :14.375   Mean   :38.44
##  3rd Qu.:18.324   3rd Qu.:43.30
##  Max.   :31.070   Max.   :96.40
```

OLS Regression

```
olsreg <- lm(Y ~ X)
summary(olsreg)
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.418  -6.388  -1.580   9.052  28.649
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  68.6190     4.7355  14.490  < 2e-16 ***
## XINC         -1.5973     0.3341  -4.780 1.83e-05 ***
## XHOVAL       -0.2739     0.1032  -2.654   0.0109 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.43 on 46 degrees of freedom
## Multiple R-squared:  0.5524, Adjusted R-squared:  0.5329
## F-statistic: 28.39 on 2 and 46 DF,  p-value: 9.341e-09
```

Spatial analysis based on contiguity

```
# Spatial weight matrix based on contiguity

listw <- nb2listw(neighbors)
summary(listw)
```

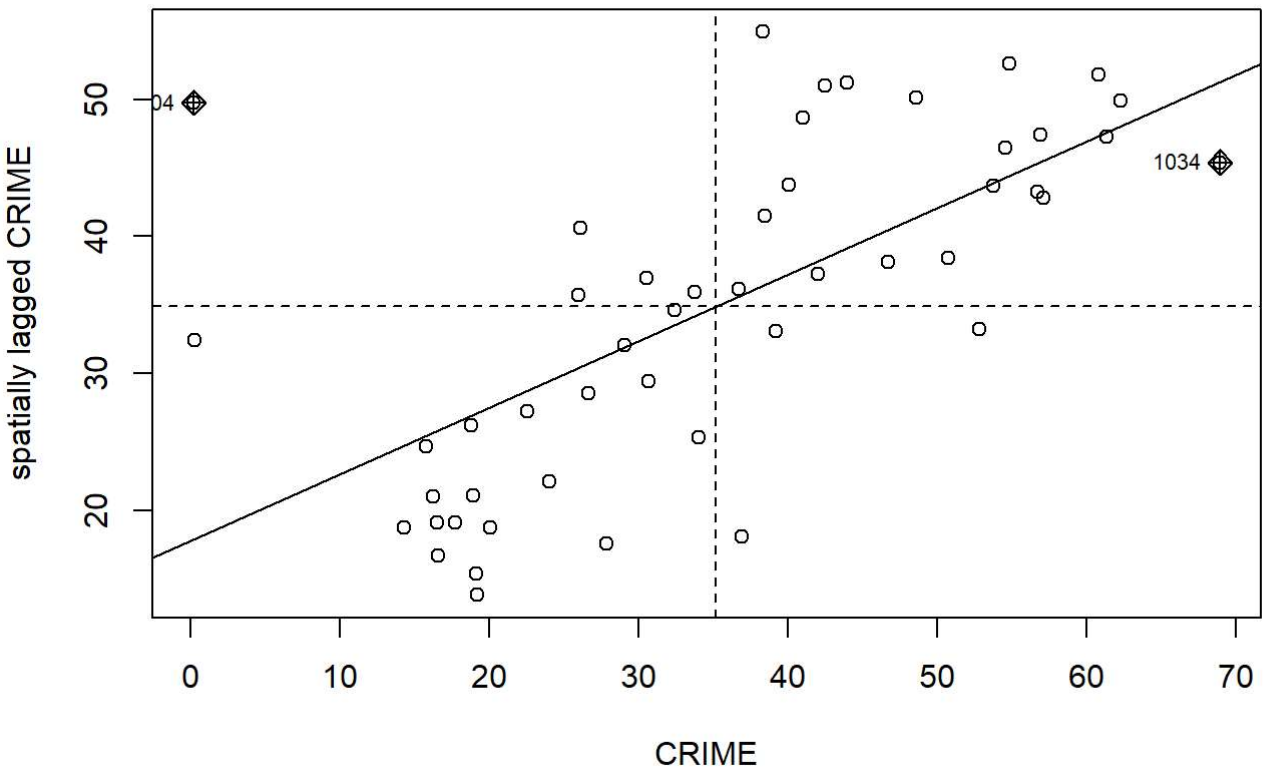
```
## Characteristics of weights list object:
## Neighbour list object:
## Number of regions: 49
## Number of nonzero links: 230
## Percentage nonzero weights: 9.579342
## Average number of links: 4.693878
## Link number distribution:
##
##  2  3  4  5  6  7  8  9 10
##  7  7 13  4  9  6  1  1  1
## 7 least connected regions:
## 1005 1008 1045 1047 1049 1048 1015 with 2 links
## 1 most connected region:
## 1017 with 10 links
##
## Weights style: W
## Weights constants summary:
##    n   nn S0       S1       S2
## W 49 2401 49 23.48489 204.6687
```

```
# Moran's I test

moran.test(CRIME, listw)
```

```
##
##  Moran I test under randomisation
##
## data:  CRIME
## weights: listw
##
## Moran I statistic standard deviate = 5.3427, p-value = 4.578e-08
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.485770914      -0.020833333       0.008991121
```

```
moran.plot(CRIME, listw)
```




```
# Lagrange multiplier test for spatial lag and spatial error dependencies
lm.LMtests(olsreg, listw, test=c("LMlag", "LMerr"))
```

```
##
##  Lagrange multiplier diagnostics for spatial dependence
##
## data:
## model: lm(formula = Y ~ X)
## weights: listw
##
## LMlag = 7.8557, df = 1, p-value = 0.005066
##
##
##  Lagrange multiplier diagnostics for spatial dependence
##
## data:
## model: lm(formula = Y ~ X)
## weights: listw
##
## LMerr = 4.6111, df = 1, p-value = 0.03177
```

```
# Spatial lag model

spatial.lag <- lagsarlm(CRIME ~ INC + HOVAL, data = mydata, listw)
summary(spatial.lag)
```

```
##
## Call:lagsarlm(formula = CRIME ~ INC + HOVAL, data = mydata, listw = listw)
##
## Residuals:
##         Min          1Q      Median          3Q         Max
## -37.4497093  -5.4565567   0.0016387   6.7159553  24.7107978
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##              Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 46.851431   7.314754  6.4051 1.503e-10
## INC         -1.073533   0.310872 -3.4533 0.0005538
## HOVAL       -0.269997   0.090128 -2.9957 0.0027381
##
## Rho: 0.40389, LR test value: 8.4179, p-value: 0.0037154
## Asymptotic standard error: 0.12071
##     z-value: 3.3459, p-value: 0.00082027
## Wald statistic: 11.195, p-value: 0.00082027
##
## Log likelihood: -183.1683 for lag model
## ML residual variance (sigma squared): 99.164, (sigma: 9.9581)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 376.34, (AIC for lm: 382.75)
## LM test for residual autocorrelation
## test value: 0.19184, p-value: 0.66139
```

```
# Spatial error model
spatial.error <- errorsarlm(CRIME ~ INC + HOVAL, data = mydata, listw)
summary(spatial.error)
```

```
##
## Call:errorsarlm(formula = CRIME ~ INC + HOVAL, data = mydata, listw = listw)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -34.45950  -6.21730  -0.69775   7.65256  24.23631
##
## Type: error
## Coefficients: (asymptotic standard errors)
##              Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 61.053618   5.314875 11.4873 < 2.2e-16
## INC         -0.995473   0.337025 -2.9537 0.0031398
## HOVAL       -0.307979   0.092584 -3.3265 0.0008794
##
## Lambda: 0.52089, LR test value: 6.4441, p-value: 0.011132
## Asymptotic standard error: 0.14129
##     z-value: 3.6868, p-value: 0.00022713
## Wald statistic: 13.592, p-value: 0.00022713
##
## Log likelihood: -184.1552 for error model
## ML residual variance (sigma squared): 99.98, (sigma: 9.999)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 378.31, (AIC for lm: 382.75)
```

Spatial analysis based on distance weight matrix

```
# Spatial weight matrix based on distance (with lower and upper bounds for distance, d1 and d2)

nb <- dnearneigh(xy, d1=0, d2=10)
listw <- nb2listw(nb, style="W")
summary(listw)
```

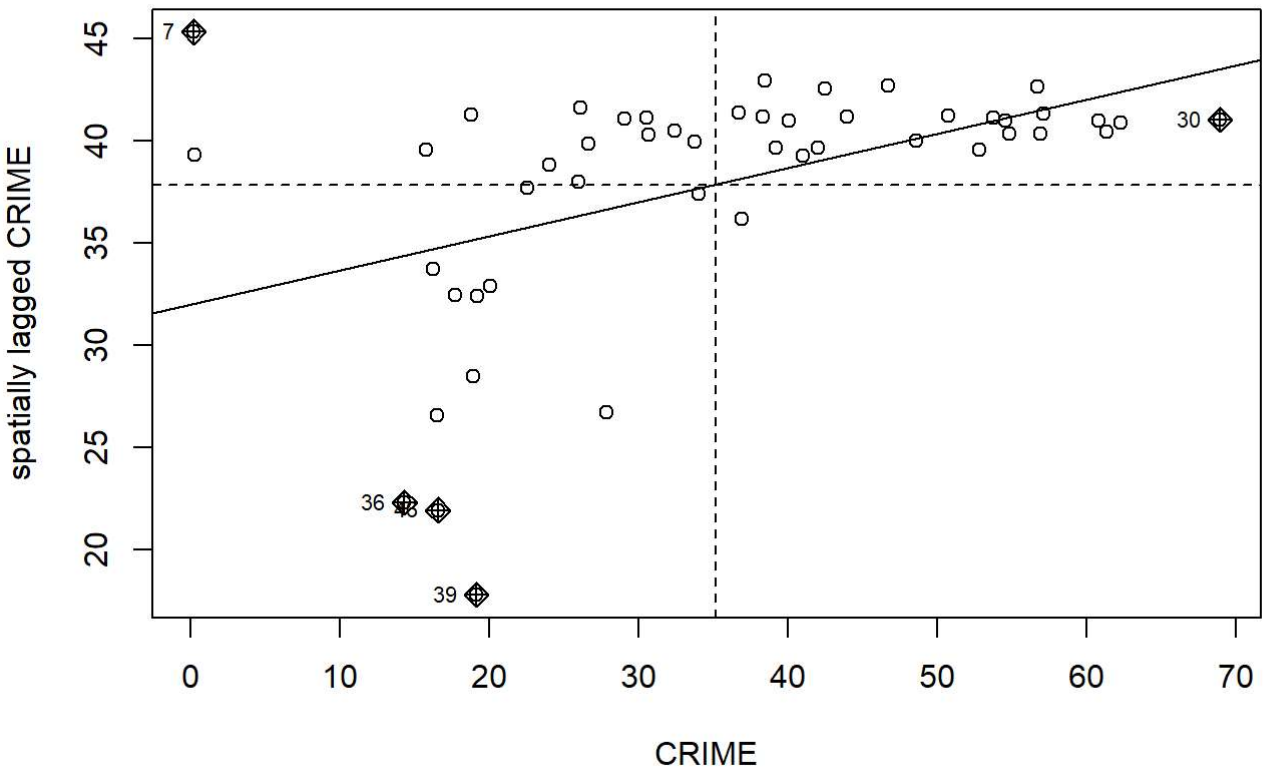
```
## Characteristics of weights list object:
## Neighbour list object:
## Number of regions: 49
## Number of nonzero links: 1234
## Percentage nonzero weights: 51.39525
## Average number of links: 25.18367
## Link number distribution:
##
##  5  6  8 10 12 13 14 15 16 18 19 20 21 24 26 28 29 30 31 32 33 34 35 36 38
##  1  2  1  1  1  1  1  1  1  1  2  2  4  3  1  2  2  2  3  3  2  6  2  2  2
## 1 least connected region:
## 39 with 5 links
## 2 most connected regions:
## 22 26 with 38 links
##
## Weights style: W
## Weights constants summary:
##    n   nn S0       S1       S2
## W 49 2401 49 4.763862 199.3227
```

```
# Moran's I test

moran.test(CRIME, listw)
```

```
##
##  Moran I test under randomisation
##
## data:  CRIME
## weights: listw
##
## Moran I statistic standard deviate = 5.6185, p-value = 9.629e-09
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.167361950      -0.020833333       0.001121945
```

```
moran.plot(CRIME, listw)
```




```
# Lagrange multiplier test for spatial lag and spatial error dependencies

lm.LMtests(olsreg, listw, test=c("LMlag", "LMerr"))
```

```
##
##  Lagrange multiplier diagnostics for spatial dependence
##
## data:
## model: lm(formula = Y ~ X)
## weights: listw
##
## LMlag = 2.8736, df = 1, p-value = 0.09004
##
##
##  Lagrange multiplier diagnostics for spatial dependence
##
## data:
## model: lm(formula = Y ~ X)
## weights: listw
##
## LMerr = 0.12837, df = 1, p-value = 0.7201
```

```
# Spatial lag model

spatial.lag1 <- lagsarlm(CRIME ~ INC + HOVAL, data = mydata, listw)
summary(spatial.lag1)
```

```
##
## Call:lagsarlm(formula = CRIME ~ INC + HOVAL, data = mydata, listw = listw)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -35.68762  -7.01341  -0.83503   8.40003  27.44614
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##              Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 51.627459  11.817848  4.3686 1.250e-05
## INC         -1.384654   0.334827 -4.1354 3.543e-05
## HOVAL       -0.281171   0.098167 -2.8642  0.004181
##
## Rho: 0.37541, LR test value: 1.9423, p-value: 0.16342
## Asymptotic standard error: 0.26881
##     z-value: 1.3966, p-value: 0.16254
## Wald statistic: 1.9504, p-value: 0.16254
##
## Log likelihood: -186.4061 for lag model
## ML residual variance (sigma squared): 117.06, (sigma: 10.82)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 382.81, (AIC for lm: 382.75)
## LM test for residual autocorrelation
## test value: 1.4922, p-value: 0.22187
```

```
# Spatial error model

spatial.error1 <- errorsarlm(CRIME ~ INC + HOVAL, data = mydata, listw)
summary(spatial.error1)
```

```
##
## Call:errorsarlm(formula = CRIME ~ INC + HOVAL, data = mydata, listw = listw)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -34.85406  -6.71581  -0.46122   8.91963  28.71276
##
## Type: error
## Coefficients: (asymptotic standard errors)
##              Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 67.311097   4.836926 13.9161 < 2.2e-16
## INC         -1.538356   0.330112 -4.6601 3.161e-06
## HOVAL       -0.268478   0.098726 -2.7194  0.006539
##
## Lambda: 0.18231, LR test value: 0.14824, p-value: 0.70022
## Asymptotic standard error: 0.4012
##     z-value: 0.45442, p-value: 0.64953
## Wald statistic: 0.20649, p-value: 0.64953
##
## Log likelihood: -187.3031 for error model
## ML residual variance (sigma squared): 122.18, (sigma: 11.054)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 384.61, (AIC for lm: 382.75)
```