Tarefa4 - GeoAnálise e Estatística Espacial

Gustavo Hotta - A56865193, Rafael Furlan - A56871487, Ricardo Squassina Lee - A56843646

Lendo o Shapefile crime\_mg:

crime\_mg<- readOGR(dsn = "crime\_mg", layer = "crime\_mg",verbose = TRUE, use\_iconv = TRUE, p4s = "+proj=longlat +ellps=WGS84")

## OGR data source with driver: ESRI Shapefile   
## Source: "C:\Users\risquass\My Tresors\zz-pessoal\FGV\git\Trabalhos\GAEE\Tarefa 4\crime\_mg", layer: "crime\_mg"  
## with 754 features  
## It has 17 fields  
## Integer64 fields read as strings: POP\_RUR POP\_URB POP\_FEM POP\_MAS

names(crime\_mg)

## [1] "CODMUNI" "ID" "MUNIC" "AREA" "INDICE94" "INDICE95"  
## [7] "GINI\_91" "POP\_94" "POP\_RUR" "POP\_URB" "POP\_FEM" "POP\_MAS"   
## [13] "POP\_TOT" "URBLEVEL" "PIB\_PC" "X\_COORD" "Y\_COORD"

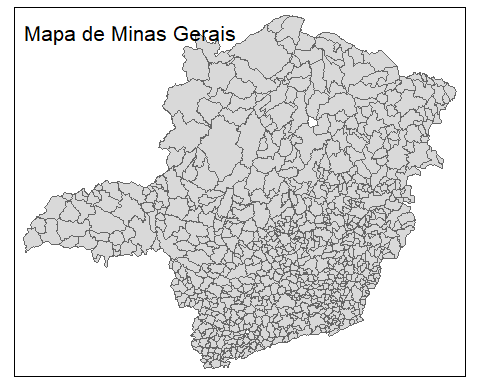
summary(crime\_mg)

## Object of class SpatialPolygonsDataFrame  
## Coordinates:  
## min max  
## x -51.06258 -39.85724  
## y -22.91696 -14.23725  
## Is projected: FALSE   
## proj4string : [+proj=longlat +ellps=WGS84]  
## Data attributes:  
## CODMUNI ID MUNIC   
## Min. : 10 Min. : 0.0 Ã\201gua Boa : 1   
## 1st Qu.:1852 1st Qu.:188.2 Ã\201gua Comprida : 1   
## Median :3645 Median :377.5 Ã\201guas Formosas : 1   
## Mean :3636 Mean :377.5 Ã\201guas Vermelhas : 1   
## 3rd Qu.:5444 3rd Qu.:566.8 AÃ§ucena : 1   
## Max. :7220 Max. :755.0 Abadia dos Dourados: 1   
## (Other) :748   
## AREA INDICE94 INDICE95 GINI\_91   
## Min. : 40.3 Min. : 0.260 Min. : 0.420 Min. :0.0000   
## 1st Qu.: 208.9 1st Qu.: 8.143 1st Qu.: 9.643 1st Qu.:0.5129   
## Median : 371.4 Median :12.060 Median :13.975 Median :0.5578   
## Mean : 779.4 Mean :13.329 Mean :15.449 Mean :0.5330   
## 3rd Qu.: 827.5 3rd Qu.:17.203 3rd Qu.:19.820 3rd Qu.:0.5960   
## Max. :13292.1 Max. :41.300 Max. :47.690 Max. :0.7127   
##   
## POP\_94 POP\_RUR POP\_URB POP\_FEM POP\_MAS   
## Min. : 820 0 : 33 0 : 33 0 : 33 0 : 33   
## 1st Qu.: 4724 1219 : 3 1374 : 3 1042 : 2 1226 : 2   
## Median : 8602 1994 : 3 10429 : 2 1070 : 2 1483 : 2   
## Mean : 21640 4995 : 3 1120 : 2 1368 : 2 1539 : 2   
## 3rd Qu.: 18054 12472 : 2 11996 : 2 1449 : 2 1643 : 2   
## Max. :2079280 1252 : 2 1257 : 2 1619 : 2 1658 : 2   
## (Other):708 (Other):710 (Other):711 (Other):711   
## POP\_TOT URBLEVEL PIB\_PC X\_COORD   
## Min. : 0 Min. :0.0000 Min. : 0 Min. :-50.81   
## 1st Qu.: 4295 1st Qu.:0.3743 1st Qu.: 1665 1st Qu.:-45.55   
## Median : 8216 Median :0.5445 Median : 2446 Median :-44.06   
## Mean : 20865 Mean :0.5373 Mean : 3036 Mean :-44.22   
## 3rd Qu.: 17710 3rd Qu.:0.7120 3rd Qu.: 3525 3rd Qu.:-42.76   
## Max. :2020161 Max. :0.9970 Max. :37728 Max. :-40.03   
##   
## Y\_COORD   
## Min. :-22.81   
## 1st Qu.:-21.19   
## Median :-20.01   
## Mean :-19.81   
## 3rd Qu.:-18.77   
## Max. :-14.46   
##

Mapa de Minas Gerais com os municípios, como no shapefile, sem tema:

tmap::qtm(crime\_mg,title = "Mapa de Minas Gerais")

## Linking to GEOS 3.6.1, GDAL 2.2.3, proj.4 4.9.3



### Pergunta 1

#### Qual das variáveis quantitativas apresentadas no shapefile crime\_mg apresenta maior auto-correlação espacial? Descreva como implementou a matriz de vizinhanca. Apresente o I de Moran e o mapa de auto-correlação espacial local (LISA map) da variável escolhida e também de pelo menos outras 3 variáveis.

##### Obs: desconsidere as variáveis Codmuni, ID, X\_coord e Y\_coord nessa analise.

Calculo de Moran’s I para verificação da auto-correlação espacial das variáveis. Aqui usamos a metodologia Rainha (queen) na matriz de vizinhança:

crime\_mg\_nb = poly2nb(crime\_mg, queen=TRUE, row.names=crime\_mg$X\_COORD)  
  
crime\_mg\_w <- nb2listw(crime\_mg\_nb, style="W")  
  
cmg\_munic <- as.numeric(crime\_mg$MUNIC)  
cmg\_area <- as.numeric(crime\_mg$AREA)  
cmg\_indice94 <- as.numeric(crime\_mg$INDICE94)  
cmg\_indice95 <- as.numeric(crime\_mg$INDICE95)  
cmg\_gini\_91 <- as.numeric(crime\_mg$GINI\_91)  
cmg\_pop\_94 <- as.numeric(crime\_mg$POP\_94)  
cmg\_pop\_rur <- as.numeric(crime\_mg$POP\_RUR)  
cmg\_pop\_urb <- as.numeric(crime\_mg$POP\_URB)  
cmg\_pop\_fem <- as.numeric(crime\_mg$POP\_FEM)  
cmg\_pop\_mas <- as.numeric(crime\_mg$POP\_MAS)  
cmg\_pop\_tot <- as.numeric(crime\_mg$POP\_TOT)  
cmg\_urblevel <- as.numeric(crime\_mg$URBLEVEL)  
cmg\_pib\_pc <- as.numeric(crime\_mg$PIB\_PC)  
  
moran\_i\_munic <- moran(cmg\_munic,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_area <- moran(cmg\_area,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_indice94 <- moran(cmg\_indice94,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_indice95 <- moran(cmg\_indice95,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_gini\_91 <- moran(cmg\_gini\_91,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_94 <- moran(cmg\_pop\_94,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_rur <- moran(cmg\_pop\_rur,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_urb <- moran(cmg\_pop\_urb,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_fem <- moran(cmg\_pop\_fem,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_mas <- moran(cmg\_pop\_mas,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pop\_tot <- moran(cmg\_pop\_tot,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_urblevel <- moran(cmg\_urblevel,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))  
moran\_i\_pib\_pc <- moran(cmg\_pib\_pc,crime\_mg\_w, length(crime\_mg\_nb), Szero(crime\_mg\_w))

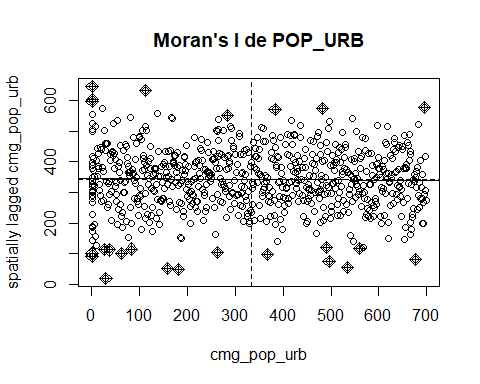
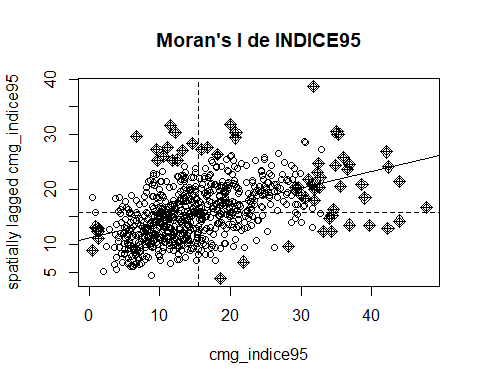
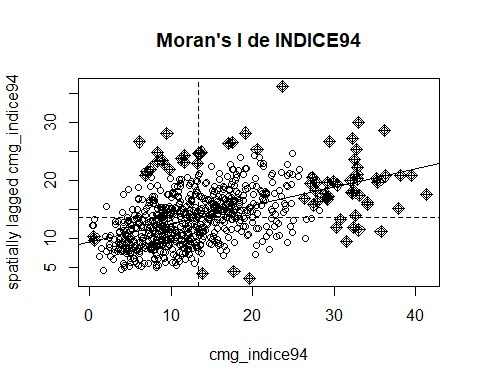
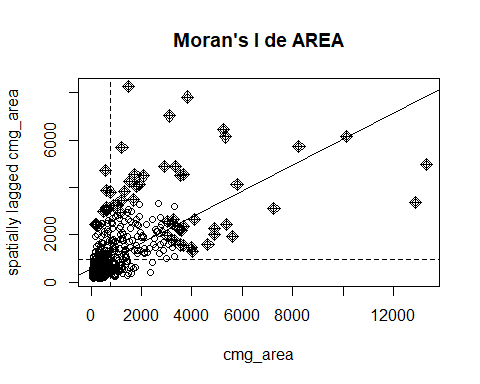
Mostrando todas as auto-correlações:

moran <- c("label","i")  
label <- c("moran\_i\_munic",  
 "moran\_i\_area",  
 "moran\_i\_indice94",  
 "moran\_i\_indice95",  
 "moran\_i\_gini\_91",  
 "moran\_i\_pop\_94",  
 "moran\_i\_pop\_rur",  
 "moran\_i\_pop\_urb",  
 "moran\_i\_pop\_fem",  
 "moran\_i\_pop\_mas",  
 "moran\_i\_pop\_tot",  
 "moran\_i\_urblevel",  
 "moran\_i\_pib\_pc"  
)  
  
moran\_i <- c(moran\_i\_munic$I  
 ,moran\_i\_area$I  
 ,moran\_i\_indice94$I  
 ,moran\_i\_indice95$I  
 ,moran\_i\_gini\_91$I  
 ,moran\_i\_pop\_94$I  
 ,moran\_i\_pop\_rur$I  
 ,moran\_i\_pop\_urb$I  
 ,moran\_i\_pop\_fem$I  
 ,moran\_i\_pop\_mas$I  
 ,moran\_i\_pop\_tot$I  
 ,moran\_i\_urblevel$I  
 ,moran\_i\_pib\_pc$I  
)  
  
moran <- data.frame(label = label, moran\_i = moran\_i)  
  
moran[order(moran$moran\_i,decreasing = TRUE),]

## label moran\_i  
## 2 moran\_i\_area 0.551407744  
## 3 moran\_i\_indice94 0.315766199  
## 4 moran\_i\_indice95 0.302517958  
## 12 moran\_i\_urblevel 0.291334697  
## 5 moran\_i\_gini\_91 0.130387445  
## 6 moran\_i\_pop\_94 0.130170365  
## 13 moran\_i\_pib\_pc 0.122220734  
## 11 moran\_i\_pop\_tot 0.121612301  
## 1 moran\_i\_munic 0.016147700  
## 9 moran\_i\_pop\_fem 0.011313607  
## 7 moran\_i\_pop\_rur 0.005987677  
## 10 moran\_i\_pop\_mas 0.004233338  
## 8 moran\_i\_pop\_urb -0.010199276

Mostrando Moran’s I das variáveis AREA, INDICE94 e INDICE95, que possuem os maiores I, e POP\_URB, que possui o menor:

{  
 moran.plot(x = cmg\_area,listw = crime\_mg\_w,labels = FALSE)  
 title("Moran's I de AREA")  
 moran.plot(x = cmg\_indice94,listw = crime\_mg\_w,labels = FALSE)  
 title("Moran's I de INDICE94")  
 moran.plot(x = cmg\_indice95,listw = crime\_mg\_w,labels = FALSE)  
 title("Moran's I de INDICE95")  
 moran.plot(x = cmg\_pop\_urb,listw = crime\_mg\_w,labels = FALSE)  
 title("Moran's I de POP\_URB")  
}



Calculando LISA

Verificando a média de links entre vizinhos:

crime\_mg\_nb

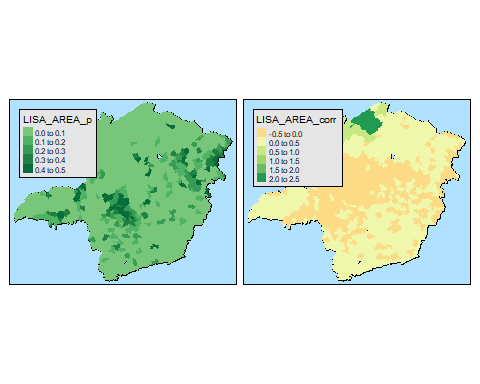
## Neighbour list object:  
## Number of regions: 754   
## Number of nonzero links: 4302   
## Percentage nonzero weights: 0.7567069   
## Average number of links: 5.70557

Como a média de links é 5.7, passamos este valor como parâmetro com o comando “mean(card(crime\_mg\_nb))” para o cálculo do LISA para as variáveis AREA - INDICE94 - INDICE95 - POP\_URB:

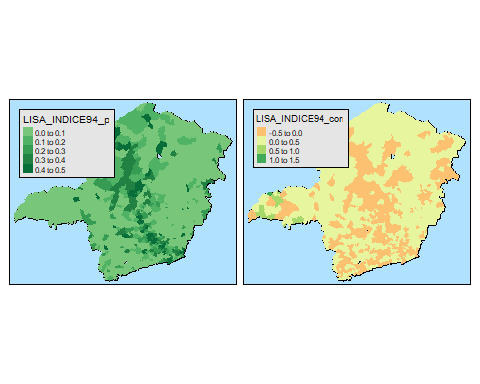
LISA\_AREA <- lisa(x = crime\_mg$X\_COORD, y = crime\_mg$Y\_COORD, z = crime\_mg$AREA, neigh = mean(card(crime\_mg\_nb)))  
LISA\_INDICE94 <- lisa(x = crime\_mg$X\_COORD, y = crime\_mg$Y\_COORD, z = crime\_mg$INDICE94, neigh = mean(card(crime\_mg\_nb)))  
LISA\_INDICE95 <- lisa(x = crime\_mg$X\_COORD, y = crime\_mg$Y\_COORD, z = crime\_mg$INDICE95, neigh = mean(card(crime\_mg\_nb)))  
LISA\_POP\_URB <- lisa(x = crime\_mg$X\_COORD, y = crime\_mg$Y\_COORD, z = as.numeric(crime\_mg$POP\_URB), neigh = mean(card(crime\_mg\_nb)))

Plot dos mapas LISA:

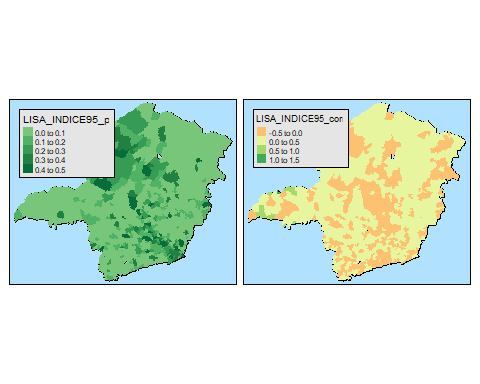
crime\_mg$LISA\_AREA\_p <- LISA\_AREA$p  
crime\_mg$LISA\_INDICE94\_p <- LISA\_INDICE94$p  
crime\_mg$LISA\_INDICE95\_p <- LISA\_INDICE95$p  
crime\_mg$LISA\_POP\_URB\_p <- LISA\_POP\_URB$p  
  
crime\_mg$LISA\_AREA\_corr <- LISA\_AREA$correlation  
crime\_mg$LISA\_INDICE94\_corr <- LISA\_INDICE94$correlation  
crime\_mg$LISA\_INDICE95\_corr <- LISA\_INDICE95$correlation  
crime\_mg$LISA\_POP\_URB\_corr <- LISA\_POP\_URB$correlation  
  
tmap::tm\_shape(crime\_mg, simplify = 1) +   
 tmap::tm\_polygons() +  
 tmap::tm\_shape(crime\_mg, simplify = 1) +  
 tmap::tm\_fill(c("LISA\_AREA\_p","LISA\_AREA\_corr"), midpoint = 0) +  
 tmap::tm\_style("natural")



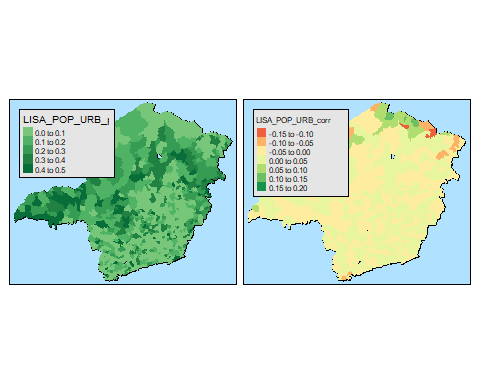
tmap::tm\_shape(crime\_mg) +  
 tmap::tm\_polygons() +  
 tmap::tm\_shape(crime\_mg, simplify = 1) +  
 tmap::tm\_fill(c("LISA\_INDICE94\_p","LISA\_INDICE94\_corr"), midpoint = 0) +   
 tmap::tm\_style("natural")



tmap::tm\_shape(crime\_mg) +  
 tmap::tm\_polygons() +  
 tmap::tm\_shape(crime\_mg, simplify = 1) +  
 tmap::tm\_fill(c("LISA\_INDICE95\_p","LISA\_INDICE95\_corr"), midpoint = 0) +   
 tmap::tm\_style("natural")

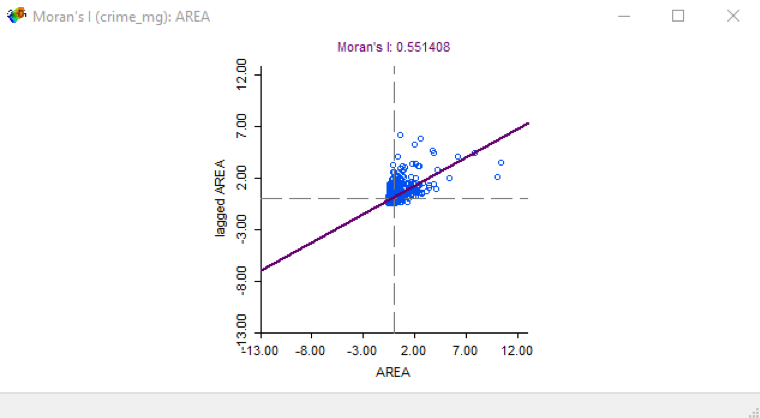


tmap::tm\_shape(crime\_mg) +  
 tmap::tm\_polygons() +  
 tmap::tm\_shape(crime\_mg, simplify = 1) +  
 tmap::tm\_fill(c("LISA\_POP\_URB\_p","LISA\_POP\_URB\_corr"), midpoint = 0) +   
 tmap::tm\_style("natural")

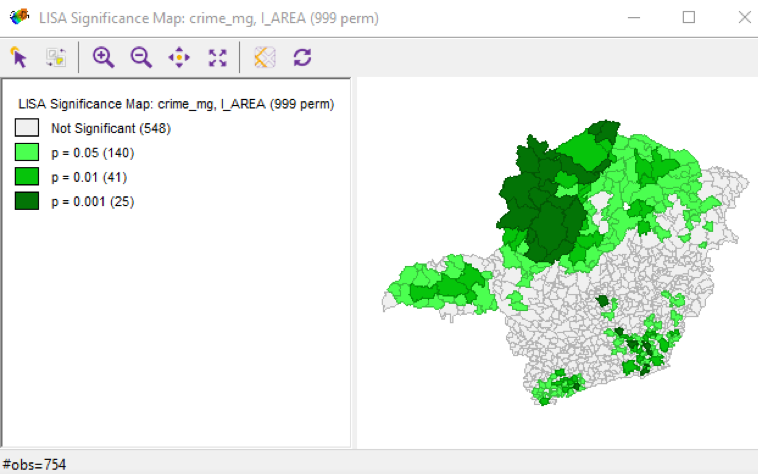


Fizemos também a implementação dos mapas LISA no GeoDa e estes foram diferentes do apresentado aqui. Abaixo, os gráficos gerados no GeoDa:

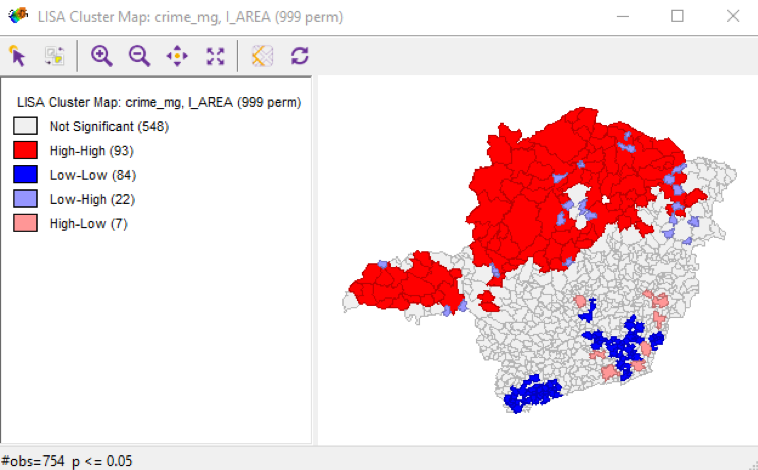
knitr::include\_graphics("MoransI-Geoda-Area.png")



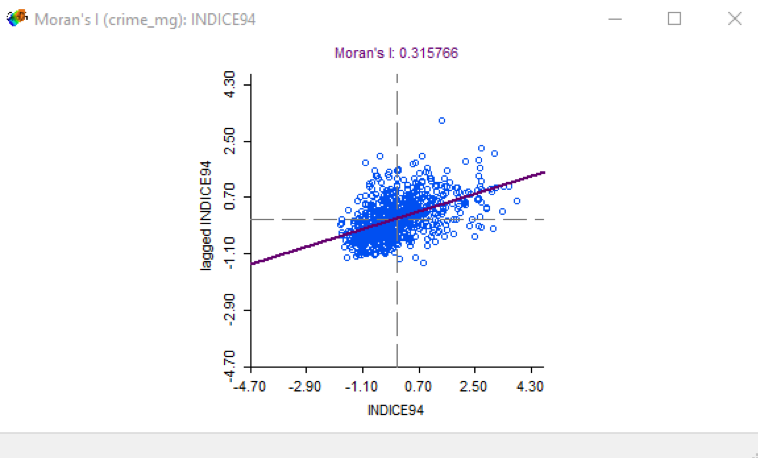
knitr::include\_graphics("LISA Significance - Geoda - Area.png")



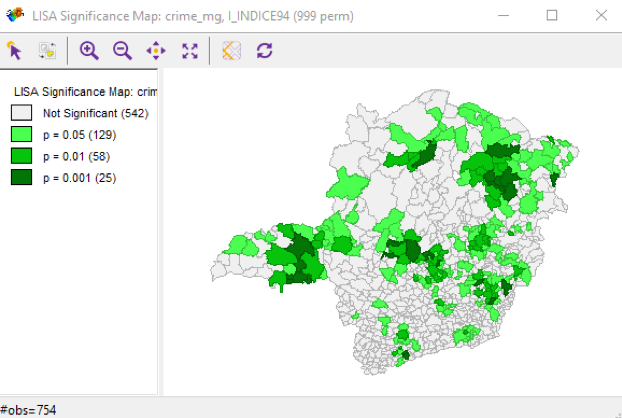
knitr::include\_graphics("LISA Cluster - Geoda - Area.png")



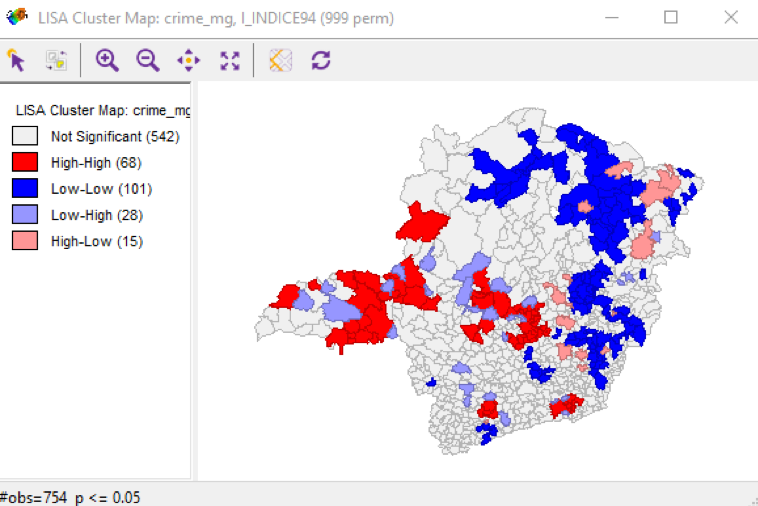
knitr::include\_graphics("MoransI-Geoda-INDICE94.png")



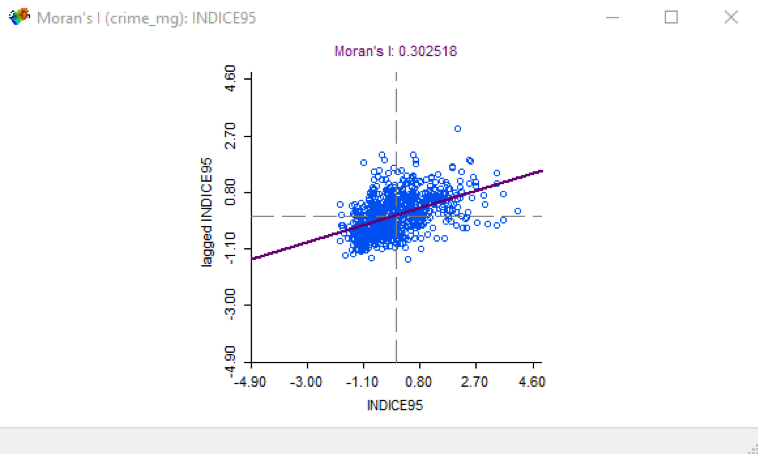
knitr::include\_graphics("LISA Significance - Geoda - INDICE94.png")



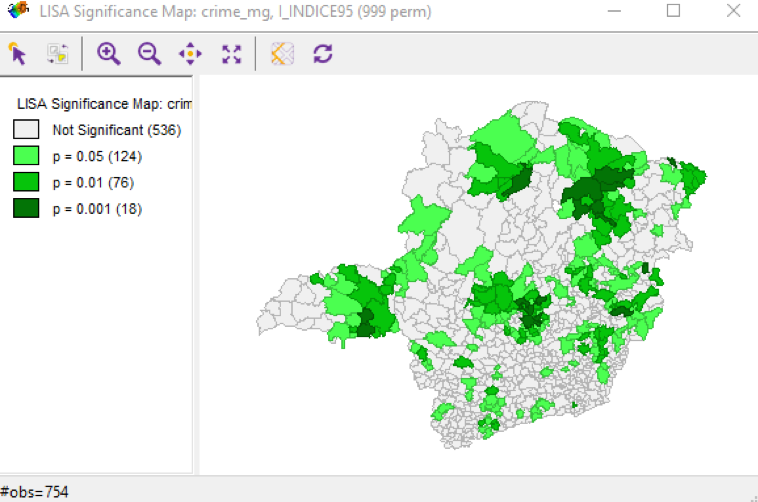
knitr::include\_graphics("LISA Cluster - Geoda - INDICE94.png")



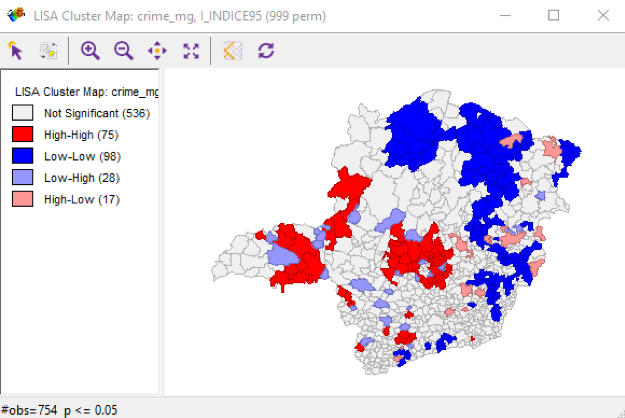
knitr::include\_graphics("MoransI-Geoda-INDICE95.png")



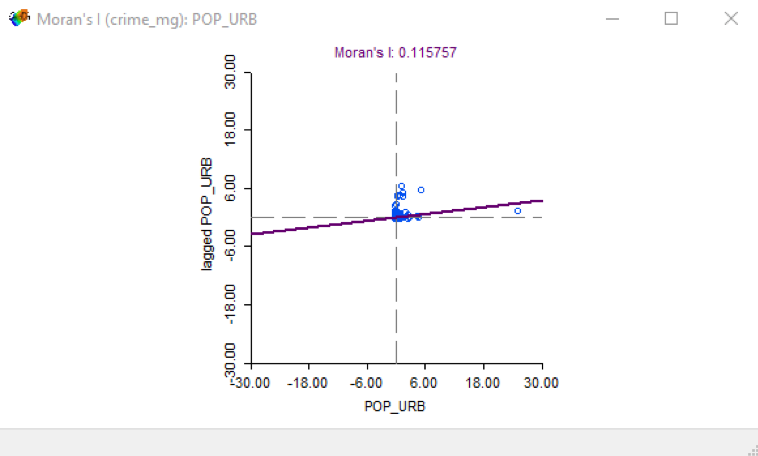
knitr::include\_graphics("LISA Significance - Geoda - INDICE95.png")



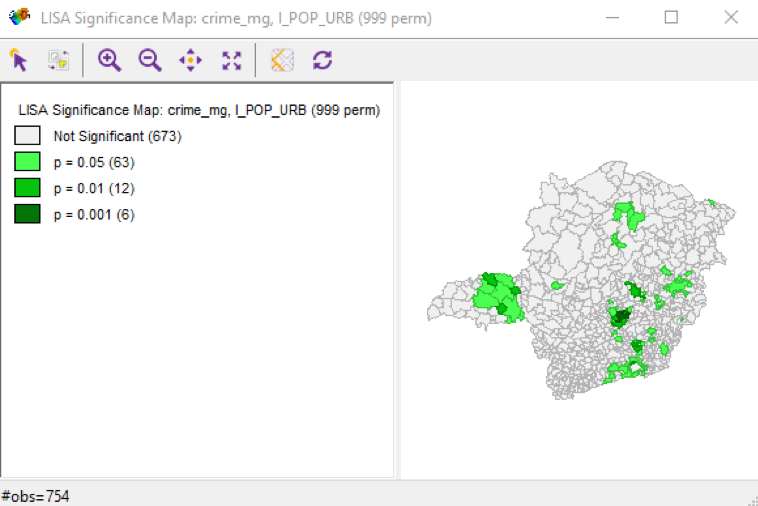
knitr::include\_graphics("LISA Cluster - Geoda - INDICE95.png")



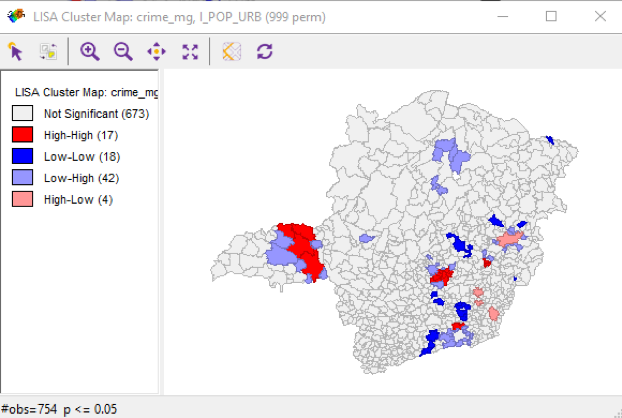
knitr::include\_graphics("MoransI-Geoda-POP\_URB.png")



knitr::include\_graphics("LISA Significance - Geoda - POP\_URB.png")



knitr::include\_graphics("LISA Cluster - Geoda - POP\_URB.png")



#### Pergunta 2

##### Implemente o modelo espacial auto-regressivo (SAR) da variável Indice95 (índice de criminalidade em 1995 de Minas Gerais) a partir de apenas uma variável independente (não pode ser Indice94, Codmuni, ID, X\_coord nem Y\_coord). Apresente o resultado da regressão linear simples e da regressão linear espacial. Apresente as equações e interprete seus coeficientes.

Regressão linear simples:

ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
head(ap)

## ID AREA INDICE94 INDICE95 GINI\_91 POP\_94 POP\_RUR POP\_URB POP\_FEM  
## 1 6 897.4 11.22 17.28 0.5372 6291 306 396 362  
## 2 7 1822.4 32.88 36.57 0.6034 21257 467 159 14  
## 3 8 636.4 4.47 7.46 0.5643 19106 67 543 681  
## 4 9 101.2 9.20 10.15 0.5069 3679 136 203 167  
## 5 5 1172.2 8.90 10.41 0.5163 24318 21 121 65  
## 6 1 1321.9 7.60 11.89 0.5796 15340 59 326 619  
## POP\_MAS POP\_TOT URBLEVEL PIB\_PC  
## 1 377 6492 0.541 3649  
## 2 9 20689 0.771 4124  
## 3 690 18961 0.304 2147  
## 4 164 3589 0.532 1720  
## 5 53 24849 0.566 1569  
## 6 624 15769 0.178 1662

lmK <- lm(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap)  
  
summary(lmK)

##   
## Call:  
## lm(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.873 -4.664 -1.174 3.639 37.569   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.3208 0.6251 10.11 <2e-16 \*\*\*  
## URBLEVEL 16.9877 1.0667 15.93 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.845 on 752 degrees of freedom  
## Multiple R-squared: 0.2522, Adjusted R-squared: 0.2512   
## F-statistic: 253.6 on 1 and 752 DF, p-value: < 2.2e-16

Regressao linear espacial - SAR

ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
crime\_mg\_nb = poly2nb(crime\_mg, queen=TRUE, row.names=crime\_mg$X\_COORD)  
crime\_mg\_w <- nb2listw(crime\_mg\_nb, style="W")  
  
sar.ap <- lagsarlm(crime\_mg$INDICE95 ~ URBLEVEL,data=ap,crime\_mg\_w,method="eigen")  
   
SARk\_SSE <- sar.ap$SSE  
  
SST <- sum((ap$INDICE95 - mean(ap$INDICE95))^2)  
  
r2\_SARk <- 1 - (SARk\_SSE/SST)  
r2\_SARk

## [1] 0.3314096

summary(sar.ap)

##   
## Call:  
## lagsarlm(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap, listw = crime\_mg\_w,   
## method = "eigen")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -15.2482 -4.2371 -1.0771 3.3952 33.9250   
##   
## Type: lag   
## Coefficients: (asymptotic standard errors)   
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 2.35307 0.79725 2.9515 0.003163  
## URBLEVEL 13.93704 1.04857 13.2914 < 2.2e-16  
##   
## Rho: 0.35437, LR test value: 66.163, p-value: 4.4409e-16  
## Asymptotic standard error: 0.044457  
## z-value: 7.9712, p-value: 1.5543e-15  
## Wald statistic: 63.539, p-value: 1.5543e-15  
##   
## Log likelihood: -2486.1 for lag model  
## ML residual variance (sigma squared): 41.776, (sigma: 6.4634)  
## Number of observations: 754   
## Number of parameters estimated: 4   
## AIC: 4980.2, (AIC for lm: 5044.4)  
## LM test for residual autocorrelation  
## test value: 10.579, p-value: 0.001144

#### Pergunta 3

##### Para essa variável que você escolheu, o modelo espacial SAR apresentou ganhos significantes com relação ao modelo linear simples? Justifique sua resposta.

##### Obs: Sugere-se fazer essa atividade no GeoDA ou no R.

cat("Rˆ2 SAR: ",r2\_SARk)

## Rˆ2 SAR: 0.3314096

cat("Rˆ2 LM:",summary(lmK)$adj.r.squared)

## Rˆ2 LM: 0.2511925

O modelo espacial SAR apresentou ganho de 8% versus o modelo linear simples.

#### Pergunta 4

##### Implemente a regressão espacial GWR da variável Indice95 (índice de criminalidade em 1995 de Minas Gerais) a partir de apenas uma variável independente (não pode ser Indice94, Codmuni, ID, X\_coord nem Y\_coord). Apresente o resultado da regressão linear simples e da regressão linear espacial por GWR. Apresente medidas da distribuição dos coeficientes (min, Q1, Q2, Q3, max), e da distribuição do R2 (min, Q1, Q2, Q3, max) e apresente os resultados globais da regressão (R2 global, basicamente).

##### Obs: Sugere-se fazer essa atividade no ArcGIS ou no R.

Regressão linear simples:

ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
head(ap)

## ID AREA INDICE94 INDICE95 GINI\_91 POP\_94 POP\_RUR POP\_URB POP\_FEM  
## 1 6 897.4 11.22 17.28 0.5372 6291 306 396 362  
## 2 7 1822.4 32.88 36.57 0.6034 21257 467 159 14  
## 3 8 636.4 4.47 7.46 0.5643 19106 67 543 681  
## 4 9 101.2 9.20 10.15 0.5069 3679 136 203 167  
## 5 5 1172.2 8.90 10.41 0.5163 24318 21 121 65  
## 6 1 1321.9 7.60 11.89 0.5796 15340 59 326 619  
## POP\_MAS POP\_TOT URBLEVEL PIB\_PC  
## 1 377 6492 0.541 3649  
## 2 9 20689 0.771 4124  
## 3 690 18961 0.304 2147  
## 4 164 3589 0.532 1720  
## 5 53 24849 0.566 1569  
## 6 624 15769 0.178 1662

lmK <- lm(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap)  
  
summary(lmK)

##   
## Call:  
## lm(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.873 -4.664 -1.174 3.639 37.569   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.3208 0.6251 10.11 <2e-16 \*\*\*  
## URBLEVEL 16.9877 1.0667 15.93 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.845 on 752 degrees of freedom  
## Multiple R-squared: 0.2522, Adjusted R-squared: 0.2512   
## F-statistic: 253.6 on 1 and 752 DF, p-value: < 2.2e-16

Regressão linear GWR:

coords <- cbind(crime\_mg$X\_COORD,crime\_mg$Y\_COORD)  
colnames(coords) <- c("X","Y")  
  
ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
bwGauss <- gwr.sel(crime\_mg$INDICE95 ~ URBLEVEL,data=ap,coords=coords,adapt=TRUE,method="aic",  
 gweight=gwr.Gauss,verbose=FALSE)  
  
gwr.ap <- gwr(crime\_mg$INDICE95 ~ URBLEVEL, data=ap,coords=coords,bandwidth=bwGauss,  
 gweight=gwr.Gauss,adapt=bwGauss,hatmatrix=TRUE)  
gwr.ap

## Call:  
## gwr(formula = crime\_mg$INDICE95 ~ URBLEVEL, data = ap, coords = coords,   
## bandwidth = bwGauss, gweight = gwr.Gauss, adapt = bwGauss,   
## hatmatrix = TRUE)  
## Kernel function: gwr.Gauss   
## Adaptive quantile: 0.013237 (about 9 of 754 data points)  
## Summary of GWR coefficient estimates at data points:  
## Min. 1st Qu. Median 3rd Qu. Max. Global  
## X.Intercept. -4.5982 4.2098 6.4983 9.3427 29.7712 6.3208  
## URBLEVEL -8.3511 10.0699 15.4027 20.0811 39.7302 16.9877  
## Number of data points: 754   
## Effective number of parameters (residual: 2traceS - traceS'S): 102.7988   
## Effective degrees of freedom (residual: 2traceS - traceS'S): 651.2012   
## Sigma (residual: 2traceS - traceS'S): 6.127271   
## Effective number of parameters (model: traceS): 71.52246   
## Effective degrees of freedom (model: traceS): 682.4775   
## Sigma (model: traceS): 5.985225   
## Sigma (ML): 5.694282   
## AICc (GWR p. 61, eq 2.33; p. 96, eq. 4.21): 4923.585   
## AIC (GWR p. 96, eq. 4.22): 4834.391   
## Residual sum of squares: 24448.34   
## Quasi-global R2: 0.4810664

GWR\_SSE <- gwr.ap$results$rss  
r2\_GWR <- 1 - (GWR\_SSE/SST)  
r2\_GWR

## [1] 0.4810664

cat("Coeficientes LM:",summary(lmK)$coefficients, "\n")

## Coeficientes LM: 6.320829 16.98768 0.6250511 1.066745 10.1125 15.92479 1.258985e-22 2.029188e-49

cat("Rˆ2 LM:",summary(lmK)$adj.r.squared, "\n")

## Rˆ2 LM: 0.2511925

cat("Coeficientes GWR:",summary(gwr.ap$lm$coefficients), "\n")

## Coeficientes GWR: 6.320829 8.987542 11.65426 11.65426 14.32097 16.98768

cat("Rˆ2 GWR:",r2\_GWR, "\n")

## Rˆ2 GWR: 0.4810664

#### Pergunta 5

##### Para essa variável que você escolheu, o modelo espacial GWR apresentou ganhos significantes com relação ao modelo linear simples? Justifique sua resposta.

cat("Rˆ2 LM:",summary(lmK)$adj.r.squared, "\n")

## Rˆ2 LM: 0.2511925

cat("Rˆ2 SAR: ",r2\_SARk, "\n")

## Rˆ2 SAR: 0.3314096

cat("Rˆ2 GWR: ",r2\_GWR, "\n")

## Rˆ2 GWR: 0.4810664

Sim, GWR Aumenta 15% o ganho em relação ao SAR, que já era 8% maior que o modelo simples.

#### Pergunta 6

##### Implemente um modelo de regressão linear multivariado stepwise da variável Indice95 (significante a 5% ou 10%, utilize o que achar melhor). Depois, “promova-o” a um modelo SAR. Apresente os resultados comparados (equaçãoo, R2). Qual modelo você escolheria como final? Se desejar, apresente mapas que sustente sua justificativa.

Implementação do modelo multivariado stepwise - regressão simples:

ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
lm.ap <- step(lm(crime\_mg$INDICE95 ~ ., data=ap))

## Start: AIC=2234.18  
## crime\_mg$INDICE95 ~ ID + AREA + INDICE94 + GINI\_91 + POP\_94 +   
## POP\_RUR + POP\_URB + POP\_FEM + POP\_MAS + POP\_TOT + URBLEVEL +   
## PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - AREA 1 3.8 14105 2232.4  
## - ID 1 4.5 14106 2232.4  
## - POP\_MAS 1 10.8 14112 2232.8  
## - POP\_FEM 1 14.6 14116 2233.0  
## - POP\_RUR 1 15.6 14117 2233.0  
## - POP\_URB 1 18.3 14119 2233.2  
## - POP\_TOT 1 18.3 14119 2233.2  
## - POP\_94 1 18.5 14120 2233.2  
## <none> 14101 2234.2  
## - PIB\_PC 1 41.7 14143 2234.4  
## - GINI\_91 1 231.7 14333 2244.5  
## - URBLEVEL 1 480.1 14581 2257.4  
## - INDICE94 1 18258.5 32360 2858.5  
##   
## Step: AIC=2232.38  
## crime\_mg$INDICE95 ~ ID + INDICE94 + GINI\_91 + POP\_94 + POP\_RUR +   
## POP\_URB + POP\_FEM + POP\_MAS + POP\_TOT + URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - ID 1 4.7 14110 2230.6  
## - POP\_MAS 1 11.9 14117 2231.0  
## - POP\_FEM 1 15.5 14120 2231.2  
## - POP\_RUR 1 16.4 14121 2231.3  
## - POP\_URB 1 18.5 14124 2231.4  
## - POP\_TOT 1 20.5 14125 2231.5  
## - POP\_94 1 20.6 14126 2231.5  
## <none> 14105 2232.4  
## - PIB\_PC 1 41.3 14146 2232.6  
## - GINI\_91 1 231.1 14336 2242.6  
## - URBLEVEL 1 477.3 14582 2255.5  
## - INDICE94 1 18441.3 32546 2860.8  
##   
## Step: AIC=2230.64  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_94 + POP\_RUR + POP\_URB +   
## POP\_FEM + POP\_MAS + POP\_TOT + URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_MAS 1 12.7 14122 2229.3  
## - POP\_RUR 1 15.9 14126 2229.5  
## - POP\_FEM 1 16.2 14126 2229.5  
## - POP\_URB 1 18.4 14128 2229.6  
## - POP\_TOT 1 19.5 14129 2229.7  
## - POP\_94 1 19.7 14129 2229.7  
## <none> 14110 2230.6  
## - PIB\_PC 1 41.1 14151 2230.8  
## - GINI\_91 1 228.9 14339 2240.8  
## - URBLEVEL 1 474.5 14584 2253.6  
## - INDICE94 1 18464.5 32574 2859.5  
##   
## Step: AIC=2229.31  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_94 + POP\_RUR + POP\_URB +   
## POP\_FEM + POP\_TOT + URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_FEM 1 3.5 14126 2227.5  
## - POP\_RUR 1 13.9 14136 2228.1  
## - POP\_TOT 1 19.2 14142 2228.3  
## - POP\_URB 1 19.4 14142 2228.3  
## - POP\_94 1 19.6 14142 2228.4  
## <none> 14122 2229.3  
## - PIB\_PC 1 40.7 14163 2229.5  
## - GINI\_91 1 230.6 14353 2239.5  
## - URBLEVEL 1 474.7 14597 2252.2  
## - INDICE94 1 18477.9 32600 2858.1  
##   
## Step: AIC=2227.5  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_94 + POP\_RUR + POP\_URB +   
## POP\_TOT + URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_URB 1 16.2 14142 2226.4  
## - POP\_TOT 1 18.9 14145 2226.5  
## - POP\_94 1 19.3 14145 2226.5  
## - POP\_RUR 1 21.3 14147 2226.6  
## <none> 14126 2227.5  
## - PIB\_PC 1 39.0 14165 2227.6  
## - GINI\_91 1 227.1 14353 2237.5  
## - URBLEVEL 1 476.7 14603 2250.5  
## - INDICE94 1 18547.7 32674 2857.8  
##   
## Step: AIC=2226.37  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_94 + POP\_RUR + POP\_TOT +   
## URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_TOT 1 18.2 14160 2225.3  
## - POP\_94 1 18.7 14161 2225.4  
## - POP\_RUR 1 21.7 14164 2225.5  
## <none> 14142 2226.4  
## - PIB\_PC 1 39.4 14181 2226.5  
## - GINI\_91 1 263.6 14406 2238.3  
## - URBLEVEL 1 465.1 14607 2248.8  
## - INDICE94 1 18809.7 32952 2862.2  
##   
## Step: AIC=2225.34  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_94 + POP\_RUR + URBLEVEL +   
## PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_94 1 2.1 14162 2223.4  
## - POP\_RUR 1 21.6 14182 2224.5  
## - PIB\_PC 1 37.4 14198 2225.3  
## <none> 14160 2225.3  
## - GINI\_91 1 257.1 14417 2236.9  
## - URBLEVEL 1 448.5 14609 2246.8  
## - INDICE94 1 18805.2 32966 2860.5  
##   
## Step: AIC=2223.45  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + POP\_RUR + URBLEVEL +   
## PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - POP\_RUR 1 21.0 14183 2222.6  
## - PIB\_PC 1 36.1 14199 2223.4  
## <none> 14162 2223.4  
## - GINI\_91 1 255.2 14418 2234.9  
## - URBLEVEL 1 447.1 14610 2244.9  
## - INDICE94 1 19048.5 33211 2864.1  
##   
## Step: AIC=2222.57  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL + PIB\_PC  
##   
## Df Sum of Sq RSS AIC  
## - PIB\_PC 1 36.5 14220 2222.5  
## <none> 14183 2222.6  
## - GINI\_91 1 234.5 14418 2232.9  
## - URBLEVEL 1 454.0 14637 2244.3  
## - INDICE94 1 19027.5 33211 2862.1  
##   
## Step: AIC=2222.5  
## crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL  
##   
## Df Sum of Sq RSS AIC  
## <none> 14220 2222.5  
## - GINI\_91 1 252.2 14472 2233.8  
## - URBLEVEL 1 550.7 14771 2249.2  
## - INDICE94 1 19206.1 33426 2864.9

lm.ap

##   
## Call:  
## lm(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
## data = ap)  
##   
## Coefficients:  
## (Intercept) INDICE94 GINI\_91 URBLEVEL   
## 4.6200 0.8295 -5.8052 5.3346

summary(lm.ap)

##   
## Call:  
## lm(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
## data = ap)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.6372 -2.4362 -0.2178 2.3819 29.1584   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.62004 0.72718 6.353 3.65e-10 \*\*\*  
## INDICE94 0.82950 0.02606 31.827 < 2e-16 \*\*\*  
## GINI\_91 -5.80519 1.59168 -3.647 0.000283 \*\*\*  
## URBLEVEL 5.33462 0.98980 5.390 9.47e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.354 on 750 degrees of freedom  
## Multiple R-squared: 0.6982, Adjusted R-squared: 0.697   
## F-statistic: 578.3 on 3 and 750 DF, p-value: < 2.2e-16

lmKmv <- lm(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
 data = ap)  
  
summary(lmKmv)

##   
## Call:  
## lm(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
## data = ap)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.6372 -2.4362 -0.2178 2.3819 29.1584   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.62004 0.72718 6.353 3.65e-10 \*\*\*  
## INDICE94 0.82950 0.02606 31.827 < 2e-16 \*\*\*  
## GINI\_91 -5.80519 1.59168 -3.647 0.000283 \*\*\*  
## URBLEVEL 5.33462 0.98980 5.390 9.47e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.354 on 750 degrees of freedom  
## Multiple R-squared: 0.6982, Adjusted R-squared: 0.697   
## F-statistic: 578.3 on 3 and 750 DF, p-value: < 2.2e-16

Implementação do modelo multivariado stepwise - regressão SAR

ap <- as.data.frame(cbind(crime\_mg$MUNIC, crime\_mg$AREA, crime\_mg$INDICE94,  
 crime\_mg$INDICE95, crime\_mg$GINI\_91, crime\_mg$POP\_94,  
 crime\_mg$POP\_RUR, crime\_mg$POP\_URB, crime\_mg$POP\_FEM,  
 crime\_mg$POP\_MAS, crime\_mg$POP\_TOT, crime\_mg$URBLEVEL,  
 crime\_mg$PIB\_PC))  
colnames(ap) <- c("ID",names(crime\_mg@data[4:15]))  
  
sarmv.ap <- lagsarlm(crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 +  
 URBLEVEL,data=ap,crime\_mg\_w,method="eigen")  
   
SARkmv\_SSE <- sarmv.ap$SSE  
  
SST <- sum((ap$INDICE95 - mean(ap$INDICE95))^2)  
  
r2\_SARkmv <- 1 - (SARkmv\_SSE/SST)  
r2\_SARkmv

## [1] 0.7029335

summary(sarmv.ap)

##   
## Call:  
## lagsarlm(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
## data = ap, listw = crime\_mg\_w, method = "eigen")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.43661 -2.38664 -0.19365 2.28926 28.43605   
##   
## Type: lag   
## Coefficients: (asymptotic standard errors)   
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 3.36490 0.82039 4.1016 4.103e-05  
## INDICE94 0.80411 0.02706 29.7163 < 2.2e-16  
## GINI\_91 -5.30895 1.58190 -3.3561 0.0007906  
## URBLEVEL 4.67297 0.99603 4.6916 2.711e-06  
##   
## Rho: 0.10647, LR test value: 10.464, p-value: 0.0012169  
## Asymptotic standard error: 0.033258  
## z-value: 3.2013, p-value: 0.0013679  
## Wald statistic: 10.249, p-value: 0.0013679  
##   
## Log likelihood: -2171.899 for lag model  
## ML residual variance (sigma squared): 18.562, (sigma: 4.3083)  
## Number of observations: 754   
## Number of parameters estimated: 6   
## AIC: 4355.8, (AIC for lm: 4364.3)  
## LM test for residual autocorrelation  
## test value: 0.10236, p-value: 0.74902

Comparação dos modelos Multi-variados de SAR e LM:

cat("Rˆ2 SAR: ",r2\_SARk, "\n")

## Rˆ2 SAR: 0.3314096

cat("Rˆ2 LM:",summary(lmK)$adj.r.squared,"\n")

## Rˆ2 LM: 0.2511925

cat("Rˆ2 SAR MV: ",r2\_SARkmv, "\n")

## Rˆ2 SAR MV: 0.7029335

cat("Rˆ2 LM MV:",summary(lmKmv)$adj.r.squared,"\n")

## Rˆ2 LM MV: 0.6969645

No modelo multivariado o ganho é de 1% sobre o modelo linear simples para o modelo SAR, porém o ganho foi de +37% com relação ao modelo original univariado.

#### Pergunta 7 (bônus)

##### Promova o modelo final linear da Pergunta 6 a um modelo GWR. Apresente os resultados comparados (equação, R2). Qual modelo você escolheria como final? Se desejar, apresente mapas que sustente sua justificativa.

Implementação do modelo multivariado stepwise - regressão GWR

bwGaussMV <- gwr.sel(crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,data=ap,coords=coords,adapt=TRUE,method="aic",  
 gweight=gwr.Gauss,verbose=FALSE)  
  
gwrMV.ap <- gwr(crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 +  
 URBLEVEL, data=ap,coords=coords,bandwidth=bwGauss,  
 gweight=gwr.Gauss,adapt=bwGaussMV,hatmatrix=TRUE)  
gwrMV.ap

## Call:  
## gwr(formula = crime\_mg$INDICE95 ~ INDICE94 + GINI\_91 + URBLEVEL,   
## data = ap, coords = coords, bandwidth = bwGauss, gweight = gwr.Gauss,   
## adapt = bwGaussMV, hatmatrix = TRUE)  
## Kernel function: gwr.Gauss   
## Adaptive quantile: 0.0411007 (about 30 of 754 data points)  
## Summary of GWR coefficient estimates at data points:  
## Min. 1st Qu. Median 3rd Qu. Max. Global  
## X.Intercept. -2.68291 2.58872 3.82700 8.52713 20.59749 4.6200  
## INDICE94 0.62641 0.75768 0.80116 0.83534 1.02374 0.8295  
## GINI\_91 -27.22302 -11.11227 -5.72057 -0.53447 13.64553 -5.8052  
## URBLEVEL -2.24401 2.25196 5.03699 8.09010 12.19130 5.3346  
## Number of data points: 754   
## Effective number of parameters (residual: 2traceS - traceS'S): 66.21814   
## Effective degrees of freedom (residual: 2traceS - traceS'S): 687.7819   
## Sigma (residual: 2traceS - traceS'S): 4.135734   
## Effective number of parameters (model: traceS): 46.62188   
## Effective degrees of freedom (model: traceS): 707.3781   
## Sigma (model: traceS): 4.078046   
## Sigma (ML): 3.949956   
## AICc (GWR p. 61, eq 2.33; p. 96, eq. 4.21): 4313.115   
## AIC (GWR p. 96, eq. 4.22): 4257.927   
## Residual sum of squares: 11764.02   
## Quasi-global R2: 0.7503001

SST <- sum((ap$INDICE95 - mean(ap$INDICE95))^2)  
   
GWR\_MV\_SSE <- gwrMV.ap$results$rss  
r2\_GWR\_MV <- 1 - (GWR\_MV\_SSE/SST)  
r2\_GWR\_MV

## [1] 0.7503001

Comparando os resultados finais:

cat("Rˆ2 GWR: ",r2\_GWR, "\n")

## Rˆ2 GWR: 0.4810664

cat("Rˆ2 SAR: ",r2\_SARk, "\n")

## Rˆ2 SAR: 0.3314096

cat("Rˆ2 LM:",summary(lmK)$adj.r.squared,"\n")

## Rˆ2 LM: 0.2511925

cat("Rˆ2 GWR MV: ",r2\_GWR\_MV, "\n")

## Rˆ2 GWR MV: 0.7503001

cat("Rˆ2 SAR MV: ",r2\_SARkmv, "\n")

## Rˆ2 SAR MV: 0.7029335

cat("Rˆ2 LM MV:",summary(lmKmv)$adj.r.squared,"\n")

## Rˆ2 LM MV: 0.6969645

Notamos que o modelo GWR tambem é beneficiado por uma análise multivariada, tendo aumentado 27%, passando de 48.1% para 75%.

#### Tarefa 8 (bônus 2)

##### Produza um mapa de alta qualidade do shapefile crime\_mg utilizando a extensão tmap. Os dois grupos que produzirem os melhores mapas ganharão 0,5 ponto adicional na nota da atividade.

##### Apresente o codigo completo e o mapa produzido em sua resposta.

ani\_map <- tmap::tm\_shape(crime\_mg, simplify = 1) +   
 tmap::tm\_fill() +  
 tmap::tm\_shape(crime\_mg) +   
 tmap::tm\_fill(c("INDICE94","INDICE95")) +  
 tmap::tm\_style(style = "natural", legend.outside = TRUE) +  
 tmap::tm\_facets(free.scales.symbol.size = FALSE, nrow=1,ncol=1) +  
 tmap::tm\_layout(main.title = "Evolucao do Indice de Criminalidade de 94-95", main.title.size = 1) +  
 tmap::tm\_polygons()  
  
tmap::tmap\_animation(ani\_map, loop = TRUE, delay=200, filename = "CRIME\_MG.gif")

knitr::include\_graphics("CRIME\_MG.gif")

