Quan_Tech_Areal_Data_Exam

Auriel Fournier

Monday, March 16, 2015

```
library(ggplot2)
library(maptools)
```

```
## Loading required package: sp
## Checking rgeos availability: TRUE
```

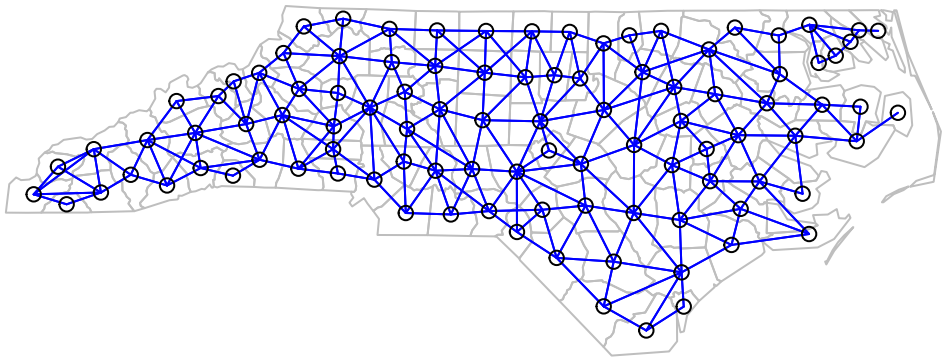
```
library(spdep)
```

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

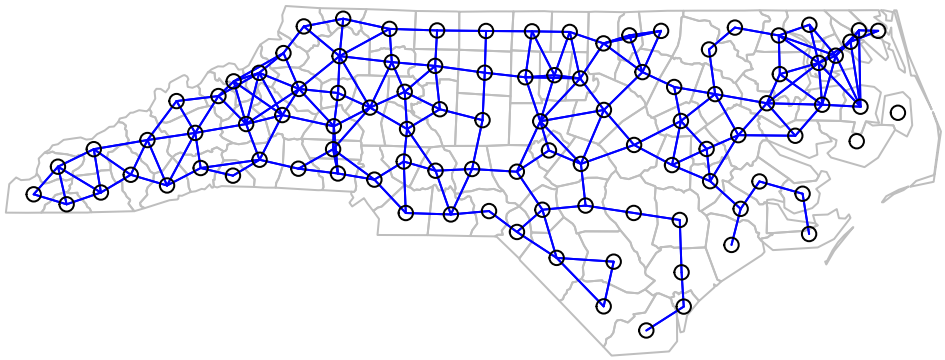
```
library(xtable)
```

```
##
## Attaching package: 'xtable'
##
## The following object is masked from 'package:maptools':
##
##     label
```

```
nc.sids <- readShapePoly(system.file("etc/shapes/sids.shp", package="spdep")[1], ID="FIPSNO", proj4stri
rn <- sapply(slot(nc.sids, "polygons"), function(x) slot(x, "ID"))
ncCC89_nb <- read.gal(system.file("etc/weights/ncCC89.gal", package="spdep")[1], region.id=rn)
ncCC85_nb <- read.gal(system.file("etc/weights/ncCR85.gal", package="spdep")[1], region.id=rn)
## Not run:
plot(nc.sids, border="grey")
plot(ncCC85_nb, coordinates(nc.sids), add=TRUE, col="blue")
```



```
plot(nc.sids, border="grey")
plot(ncCC89_nb, coordinates(nc.sids), add=TRUE, col="blue")
```



```
sids.ft.all.79 = sqrt(1000)*(sqrt(nc.sids$SID79/nc.sids$BIR79) +sqrt((nc.sids$SID79+1)/nc.sids$BIR79))

nwbirth.ft.all.79 = sqrt(1000)*(sqrt(nc.sids$NWBIR79/nc.sids$BIR79)+sqrt((nc.sids$NWBIR79+1)/nc.sids$BI

sids.ft.all.74 = sqrt(1000)*(sqrt(nc.sids$SID74/nc.sids$BIR74) +sqrt((nc.sids$SID74+1)/nc.sids$BIR74))

nwbirth.ft.all.74 = sqrt(1000)*(sqrt(nc.sids$NWBIR74/nc.sids$BIR74)+sqrt((nc.sids$NWBIR74+1)/nc.sids$BI
```

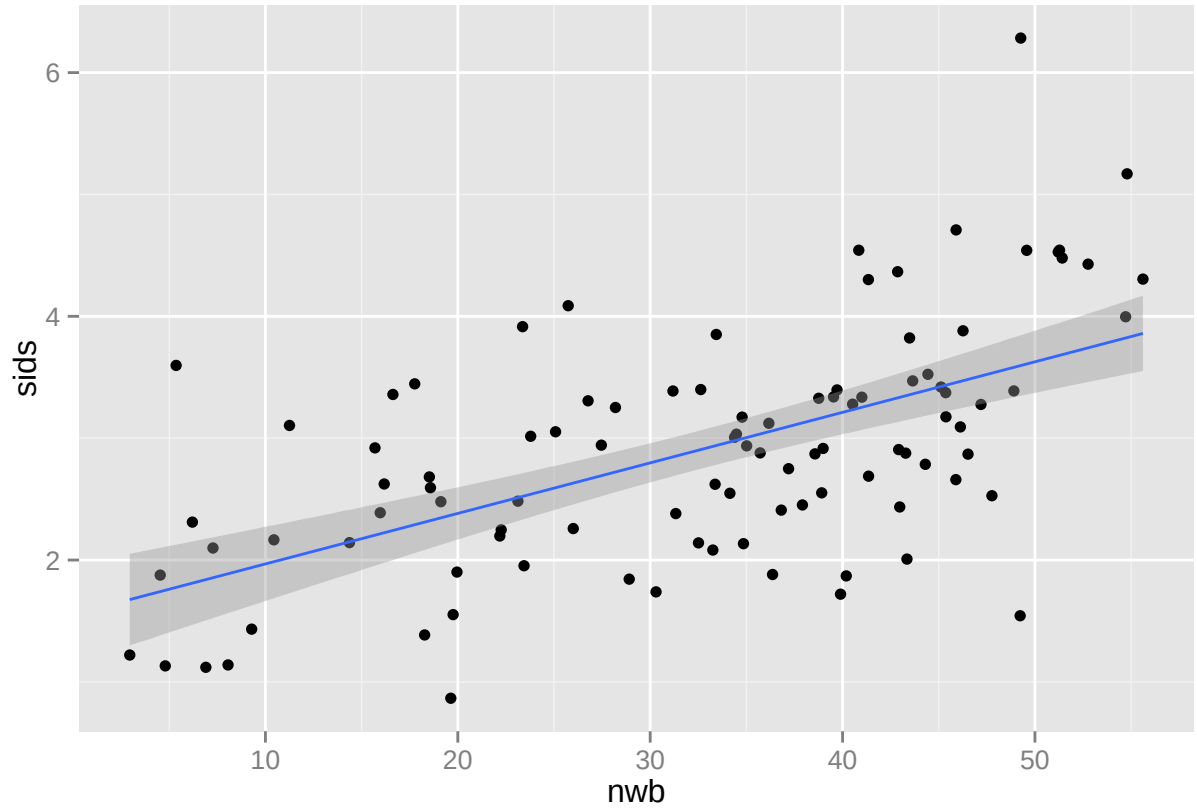
First I'd like to examine any relationship between these two variables without taking into account spatial variation. I'll do this by running a linear model of non-white births vs sids.

```
dat74 <- as.data.frame(cbind(nwbirth.ft.all.74, sids.ft.all.74))
colnames(dat74) <- c("nwb","sids")

l74 <- lm(nwb ~ sids, data=dat74)
summary(l74)
```

```
##
## Call:
## lm(formula = nwb ~ sids, data = dat74)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -32.975  -8.848   3.054   7.730  27.899
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    8.564      3.543   2.417   0.0175 *
## sids           8.275      1.157   7.155 1.53e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.23 on 98 degrees of freedom
## Multiple R-squared:  0.3431, Adjusted R-squared:  0.3364
## F-statistic: 51.19 on 1 and 98 DF,  p-value: 1.533e-10
```

```
ggplot(data=dat74,aes(nwb, sids))+geom_point()+geom_smooth(method=lm)
```

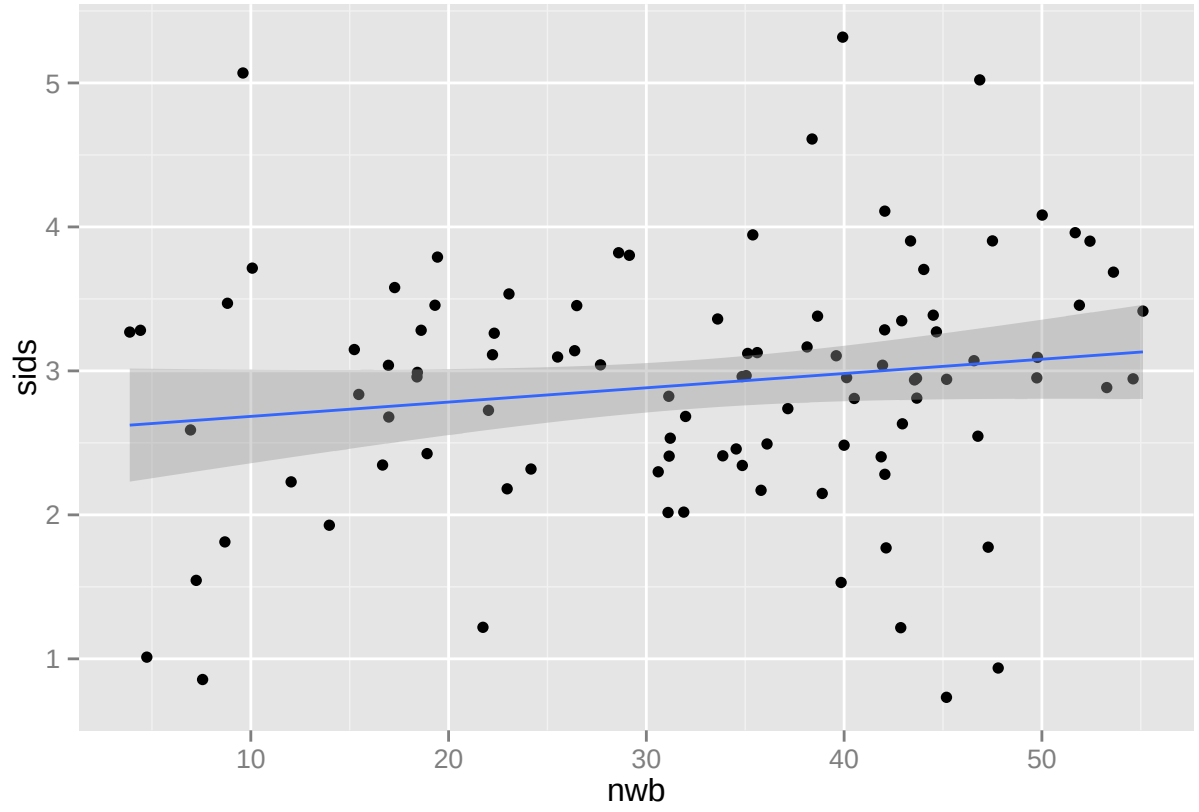



```
dat79 <- as.data.frame(cbind(nwbirth.ft.all.79, sids.ft.all.79))
colnames(dat79) <- c("nwb","sids")

l79 <- lm(nwb ~ sids, data=dat79)
summary(l79)
```

```
##
## Call:
## lm(formula = nwb ~ sids, data = dat79)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -29.547 -11.057   2.296  10.874  22.012
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   25.123      4.833   5.198  1.1e-06 ***
## sids           2.539      1.595   1.592    0.115
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.62 on 98 degrees of freedom
## Multiple R-squared:  0.02521,    Adjusted R-squared:  0.01526
## F-statistic: 2.534 on 1 and 98 DF,  p-value: 0.1146
```

```
ggplot(data=dat79,aes(nwb, sids))+geom_point()+geom_smooth(method=lm)
```



So in 1974 there is a significant relationship between the two, but not in 1979. In both cases the R is low (1974 it is .3431) suggesting that even when it is significant there might be other factors at play.

So lets start to examine things spatially.

```
col.qb <- nb2listw(poly2nb(nc.sids, queen=T),style="B")
col.rw <- nb2listw(poly2nb(nc.sids, queen=F),style="W")
col.rb <- nb2listw(poly2nb(nc.sids, queen=F),style="B")
coords <- coordinates(nc.sids)
nc.d2 <- nb2listw(dnearneigh(coords, d1=0, d=2))
nc.d1 <- nb2listw(dnearneigh(coords, d1=0, d=1))
```

```
knn3 <- nb2listw(knn2nb(knearneigh(coords, k=3)))
knn6 <- nb2listw(knn2nb(knearneigh(coords, k=6)))
knn9 <- nb2listw(knn2nb(knearneigh(coords, k=9)))
lm.morantest(174, col.qb, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: col.qb
##
## Moran I statistic standard deviate = 8.2584, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.495500417       -0.012468421        0.003783365
```

```
lm.morantest(174, col.rw, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: col.rw
##
## Moran I statistic standard deviate = 8.5295, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.553867655       -0.012762206        0.004413188
```

```
lm.morantest(174, col.rb, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: col.rb
##
## Moran I statistic standard deviate = 7.9767, p-value = 7.513e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##         0.49340320        -0.01273419         0.00402613
```

```
lm.morantest(174, nc.d2, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals
```

```
##
```

```
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: nc.d2
##
## Moran I statistic standard deviate = 26.5026, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##       0.4295462606      -0.0110250727       0.0002763484
```

```
lm.morantest(174, nc.d1, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals)
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: nc.d1
##
## Moran I statistic standard deviate = 15.3879, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.510443317       -0.011915247        0.001152339
```

```
lm.morantest(174, knn3, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals)
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: knn3
##
## Moran I statistic standard deviate = 7.1439, p-value = 4.537e-13
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.535031234       -0.012721754        0.005878984
```

```
lm.morantest(174, knn6, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals)
```

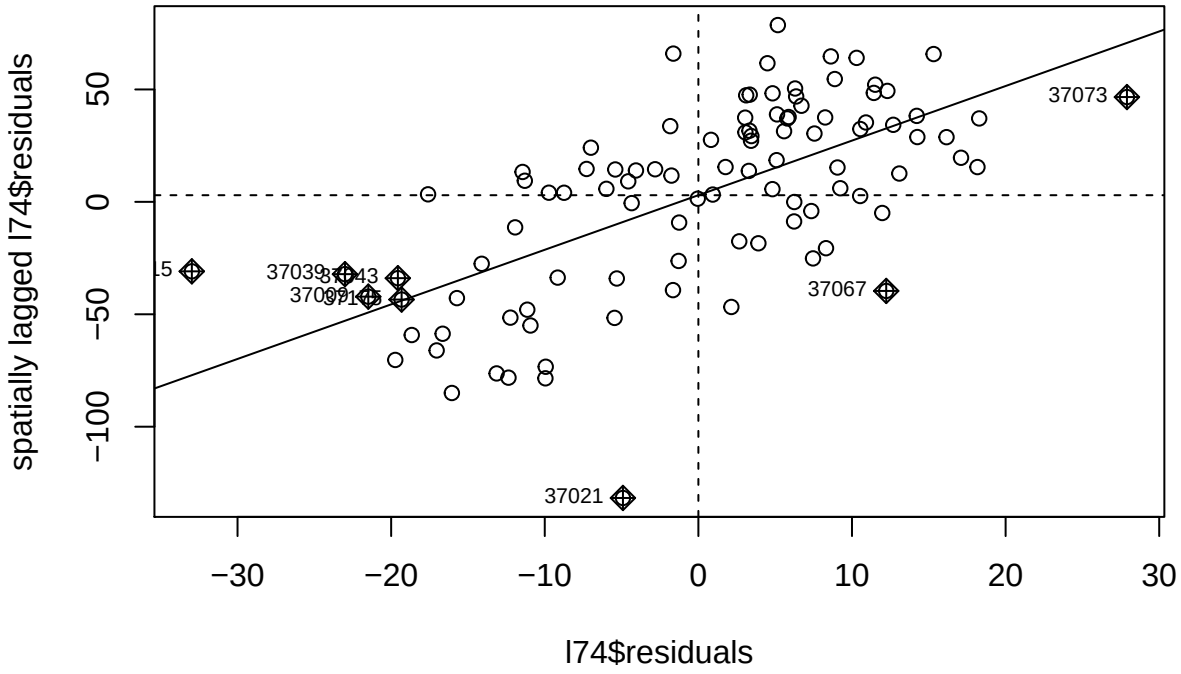
```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: knn6
##
## Moran I statistic standard deviate = 10.0734, p-value < 2.2e-16
```

```
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.521270083       -0.012398079        0.002806666
```

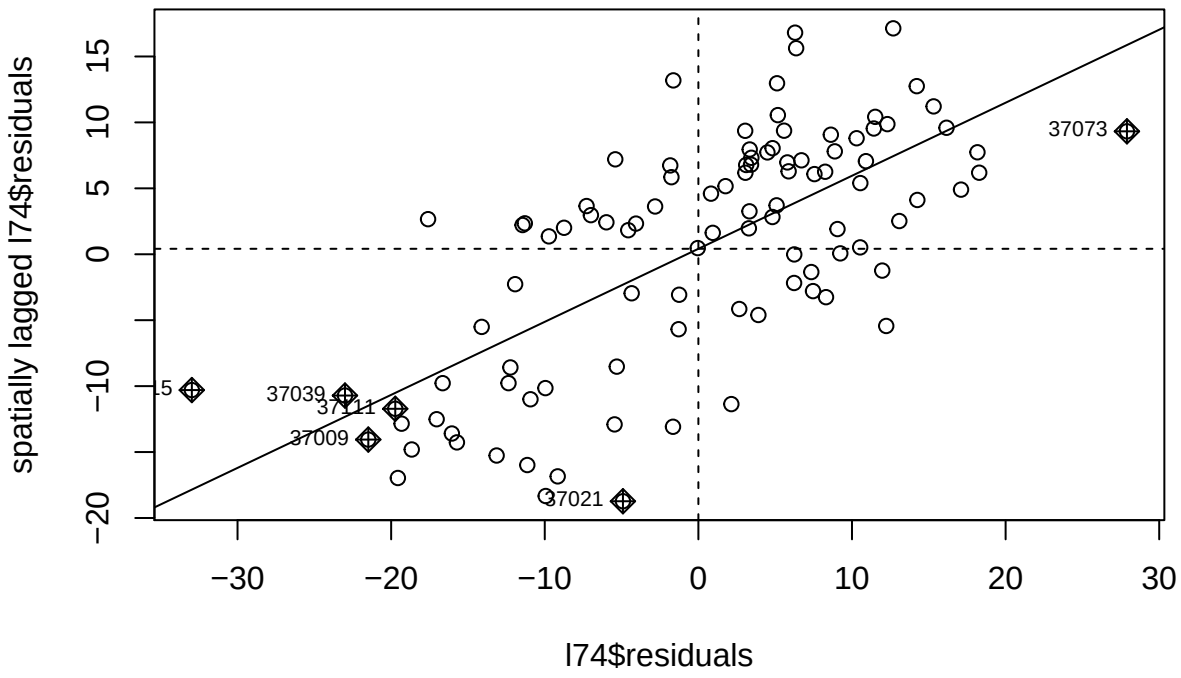
```
lm.morantest(l74, knn9, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weighted.residuals)
```

```
##
##  Global Moran's I for regression residuals
##
## data:
## model: lm(formula = nwb ~ sids, data = dat74)
## weights: knn9
##
## Moran I statistic standard deviate = 12.7183, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Observed Moran's I        Expectation           Variance
##        0.526419584       -0.012050594        0.001792528
```

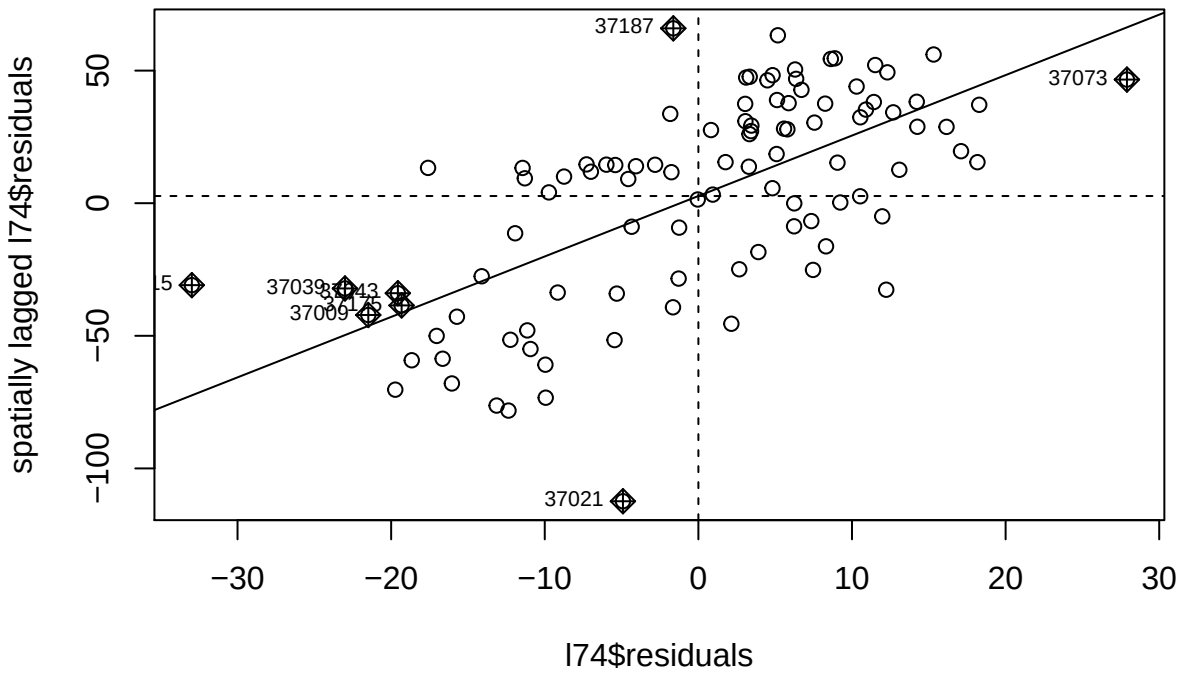
```
moran.plot(l74$residuals, listw=col.qb)
```



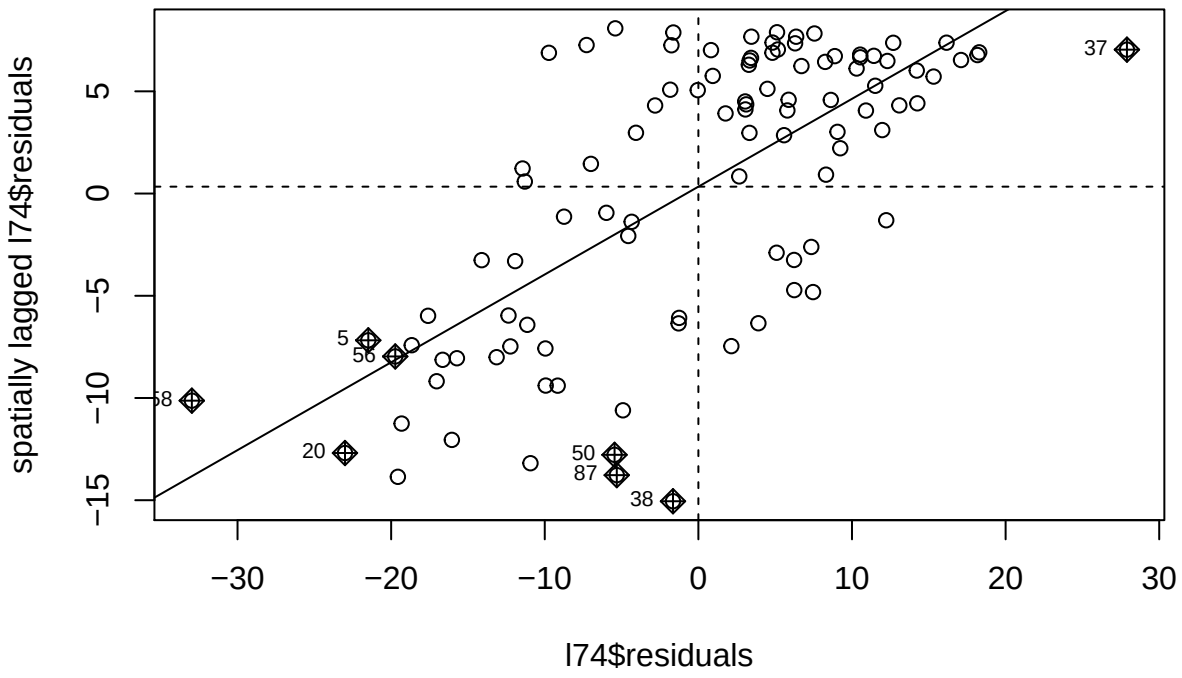

```
moran.plot(l74$residuals, listw=col.rw)
```



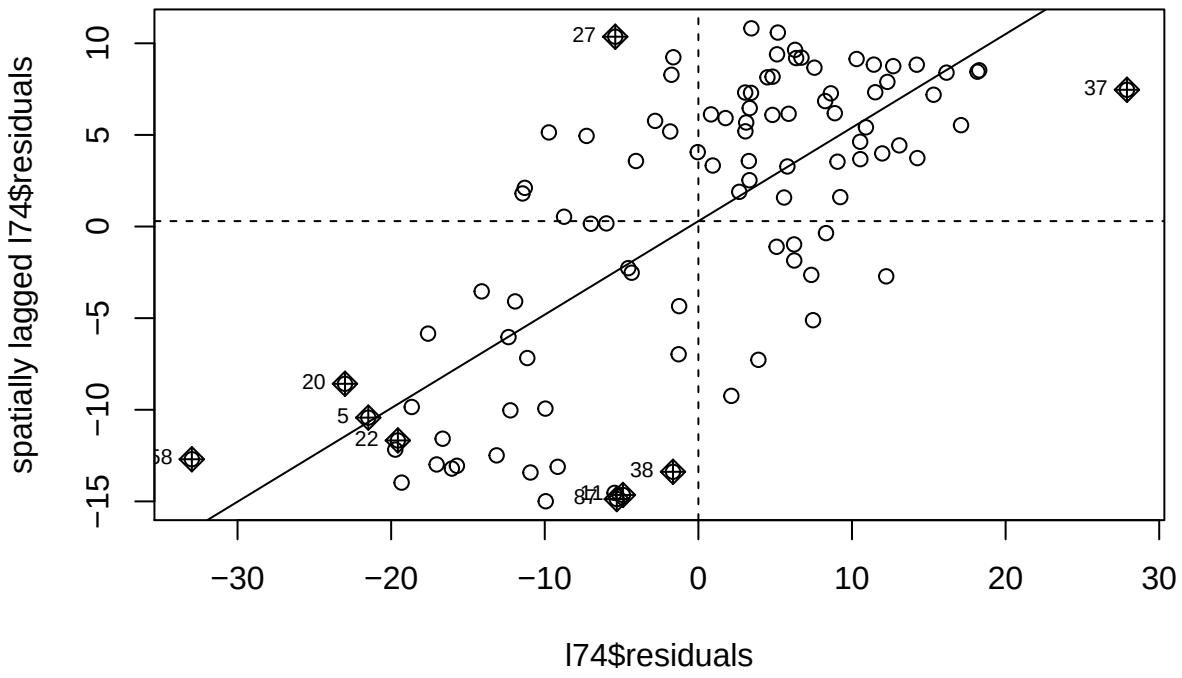

```
moran.plot(l74$residuals, listw=col.rb)
```



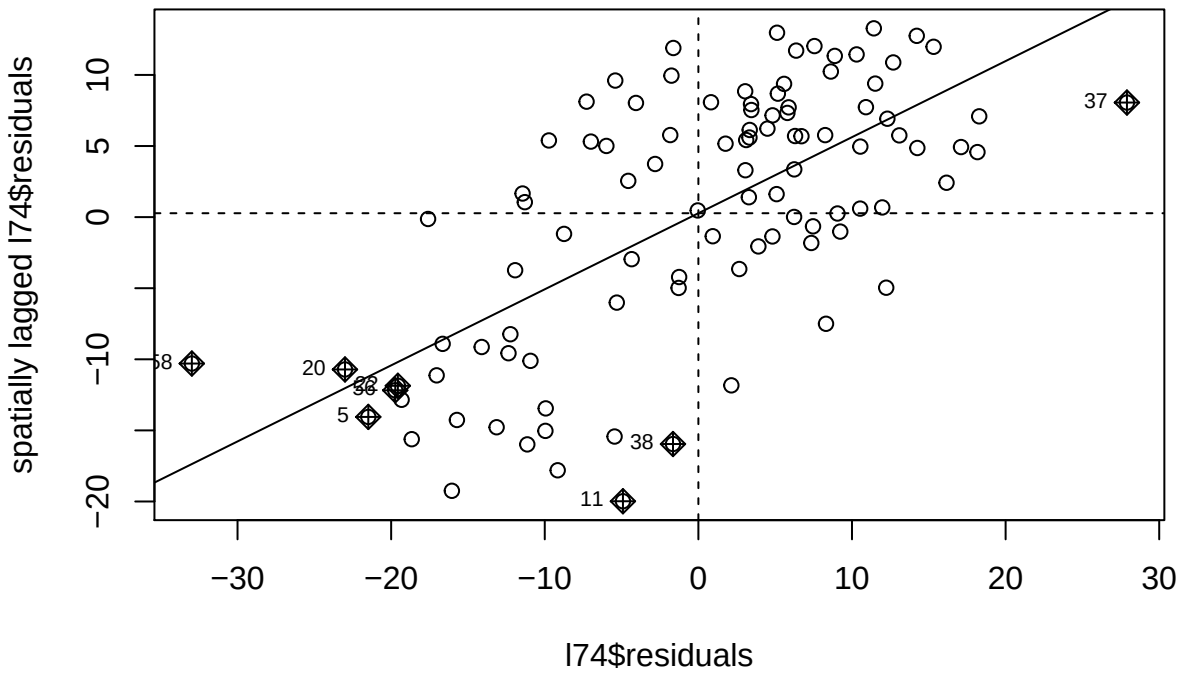

```
moran.plot(l74$residuals, listw=nc.d2)
```



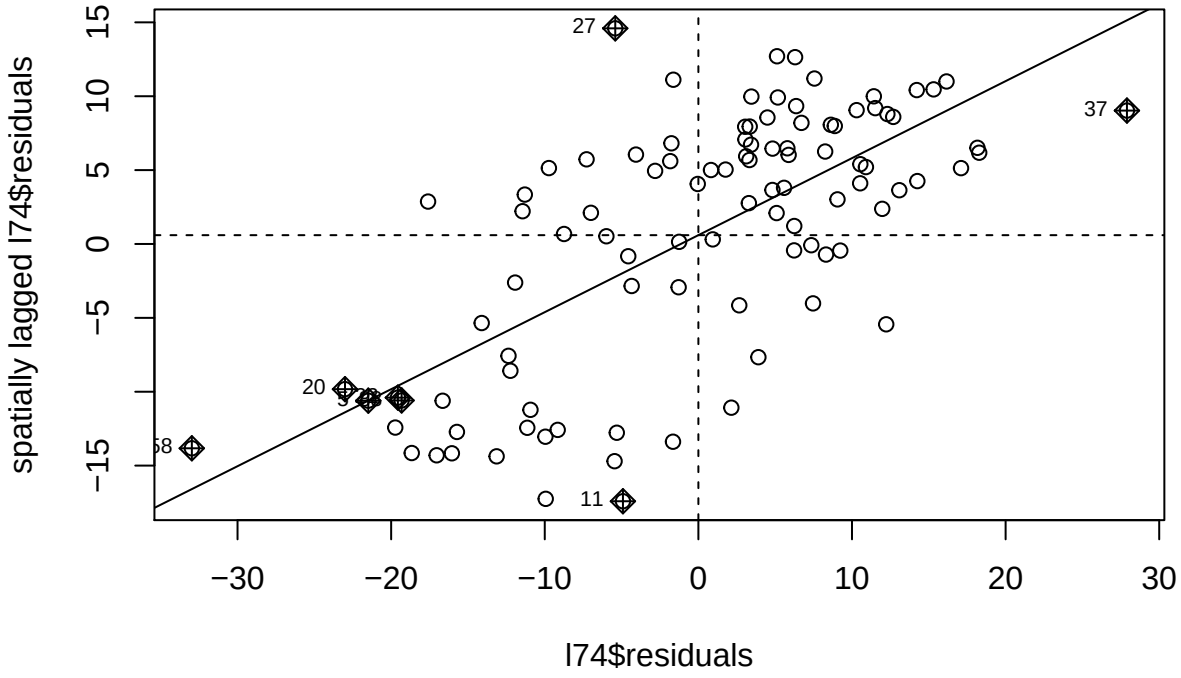

```
moran.plot(l74$residuals, listw=nc.d1)
```



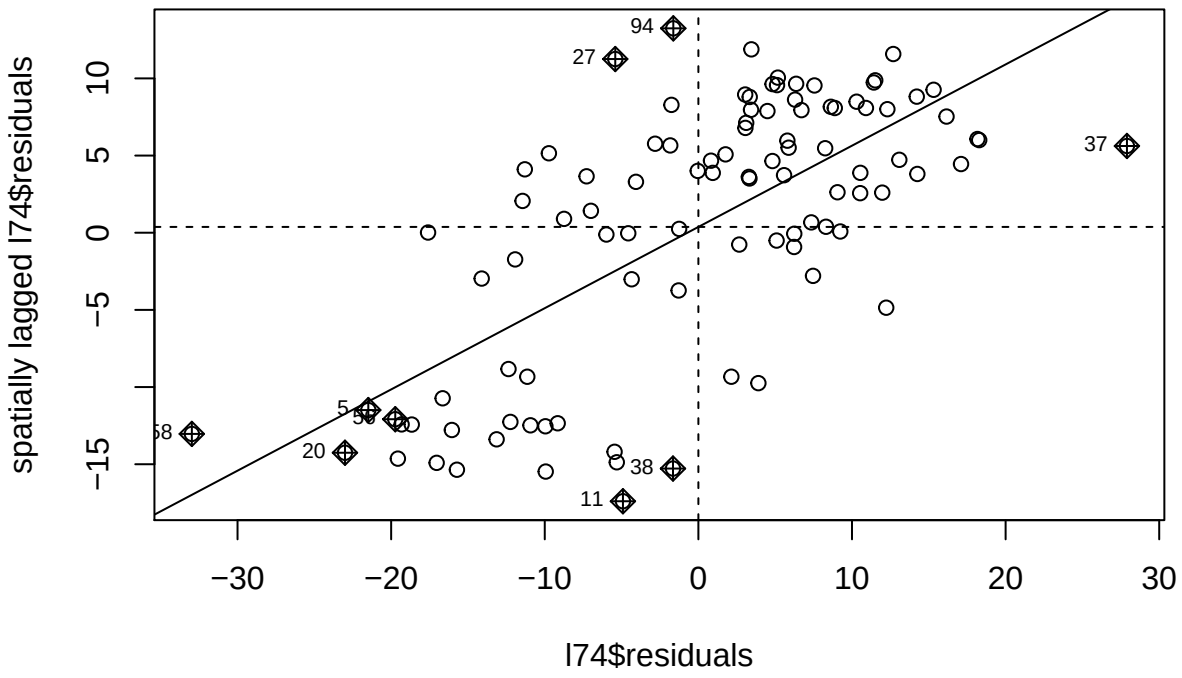

```
moran.plot(l74$residuals, listw=knn3)
```




```
moran.plot(l74$residuals, listw=knn6)
```




```
moran.plot(l74$residuals, listw=knn9)
```

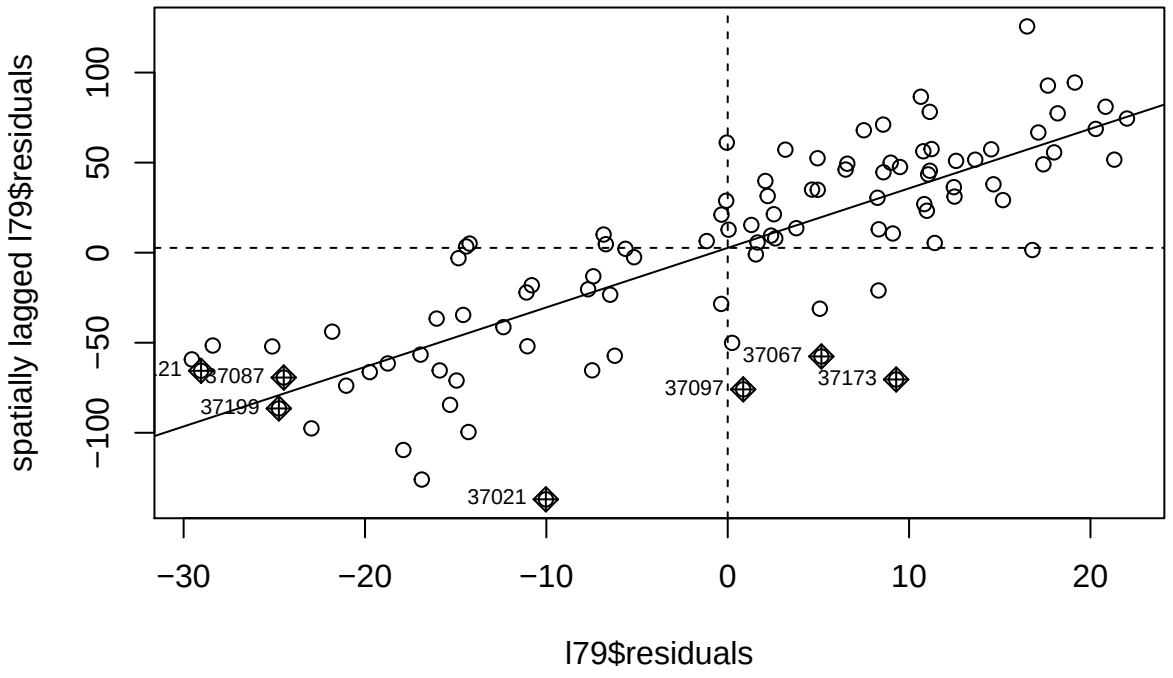



```
col.qb <- nb2listw(poly2nb(nc.sids, queen=T),style="B")
col.rw <- nb2listw(poly2nb(nc.sids, queen=F),style="W")
col.rb <- nb2listw(poly2nb(nc.sids, queen=F),style="B")
coords <- coordinates(nc.sids)
nc.d2 <- nb2listw(dnearneigh(coords, d1=0, d=2))
nc.d1 <- nb2listw(dnearneigh(coords, d1=0, d=1))
knn1 <- nb2listw(knn2nb(knearneigh(coords, k=1)))
knn2 <- nb2listw(knn2nb(knearneigh(coords, k=2)))
knn3 <- nb2listw(knn2nb(knearneigh(coords, k=3)))
knn4 <- nb2listw(knn2nb(knearneigh(coords, k=4)))
knn5 <- nb2listw(knn2nb(knearneigh(coords, k=5)))
knn6 <- nb2listw(knn2nb(knearneigh(coords, k=6)))
knn7 <- nb2listw(knn2nb(knearneigh(coords, k=7)))
knn8 <- nb2listw(knn2nb(knearneigh(coords, k=8)))
knn9 <- nb2listw(knn2nb(knearneigh(coords, k=9)))
knn10 <- nb2listw(knn2nb(knearneigh(coords, k=10)))
moranIs <- as.data.frame(matrix(ncol=3,nrow=8))
moranIs[1,] <- lm.morantest(l79, col.qb, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=wei
moranIs[2,] <-lm.morantest(l79, col.rw, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weig
moranIs[3,] <-lm.morantest(l79, col.rb, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weig
moranIs[4,] <-lm.morantest(l79, nc.d2, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weigh
moranIs[5,] <-lm.morantest(l79, nc.d1, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weigh
moranIs[6,] <-lm.morantest(l79, knn3, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weight
moranIs[7,] <-lm.morantest(l79, knn6, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weight
moranIs[8,] <-lm.morantest(l79, knn9, zero.policy=NULL, alternative="greater", spChk=NULL, resfun=weight
```

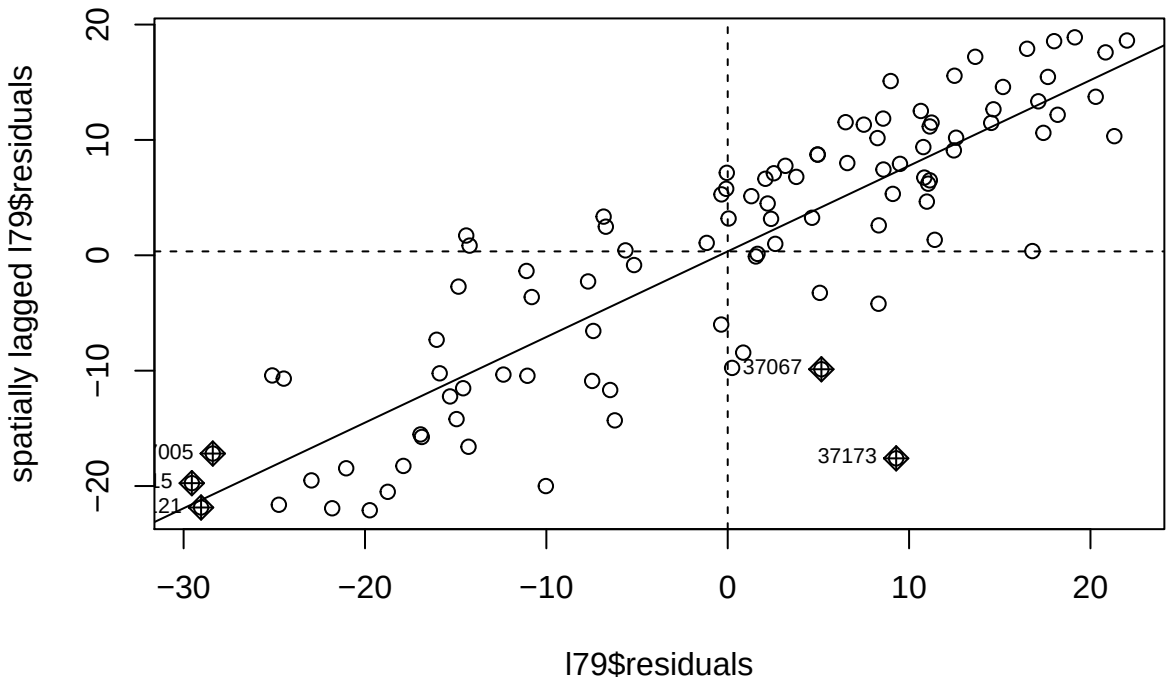
```
names <- c("col.qb","col.rw","col.rb","nc.d2","nd.d1","knn3","knn6","knn9")
moranIs$names <- names

colnames(moranIs) <- c("Observed Moran's I","Expectation","Variance","Weighting Scheme")
```

Observed Moran's I	Expectation	Variance	Weighting Scheme
0.675	-0.011	0.004	col.qb
0.742	-0.012	0.004	col.rw
0.682	-0.012	0.004	col.rb
0.527	-0.010	0.000	nc.d2
0.634	-0.010	0.001	nd.d1
0.740	-0.011	0.006	knn3
0.704	-0.011	0.003	knn6
0.694	-0.010	0.002	knn9

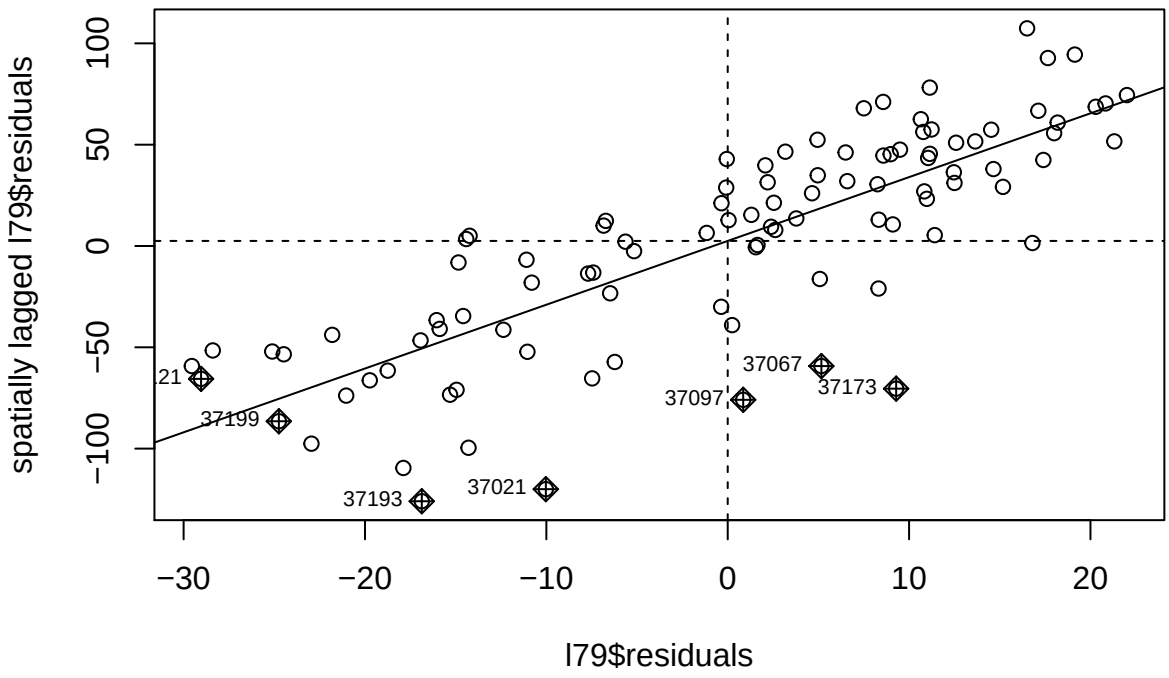
```
moran.plot(l79$residuals, listw=col.qb)
```



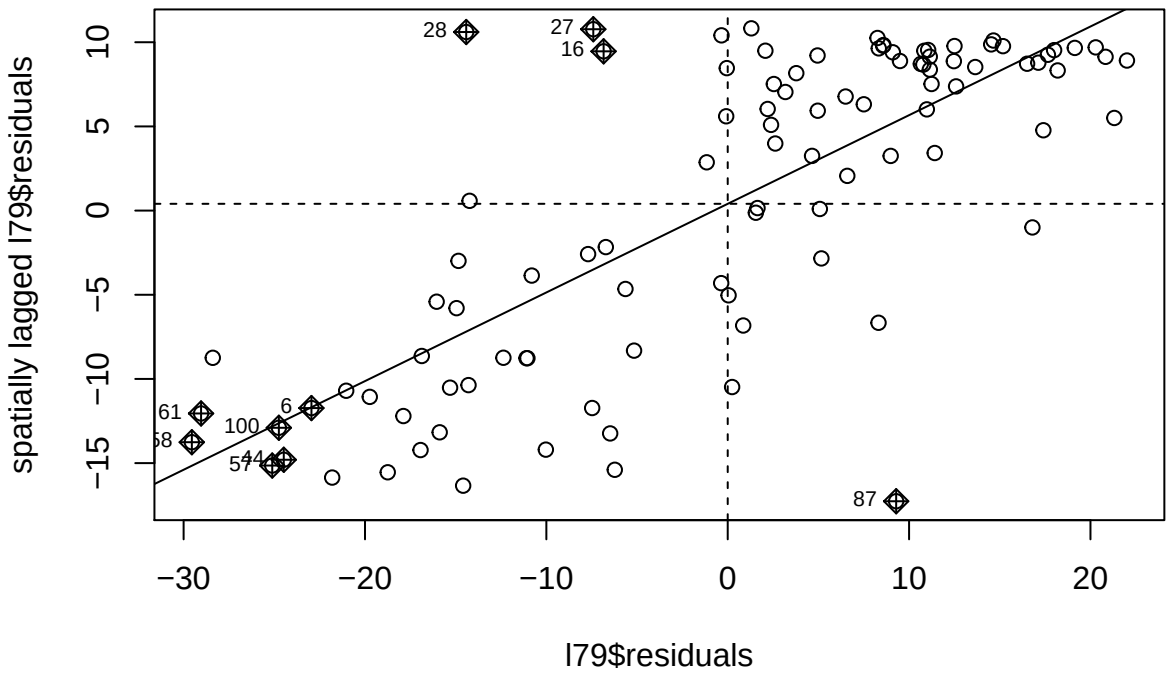

```
moran.plot(l79$residuals, listw=col.rw)
```



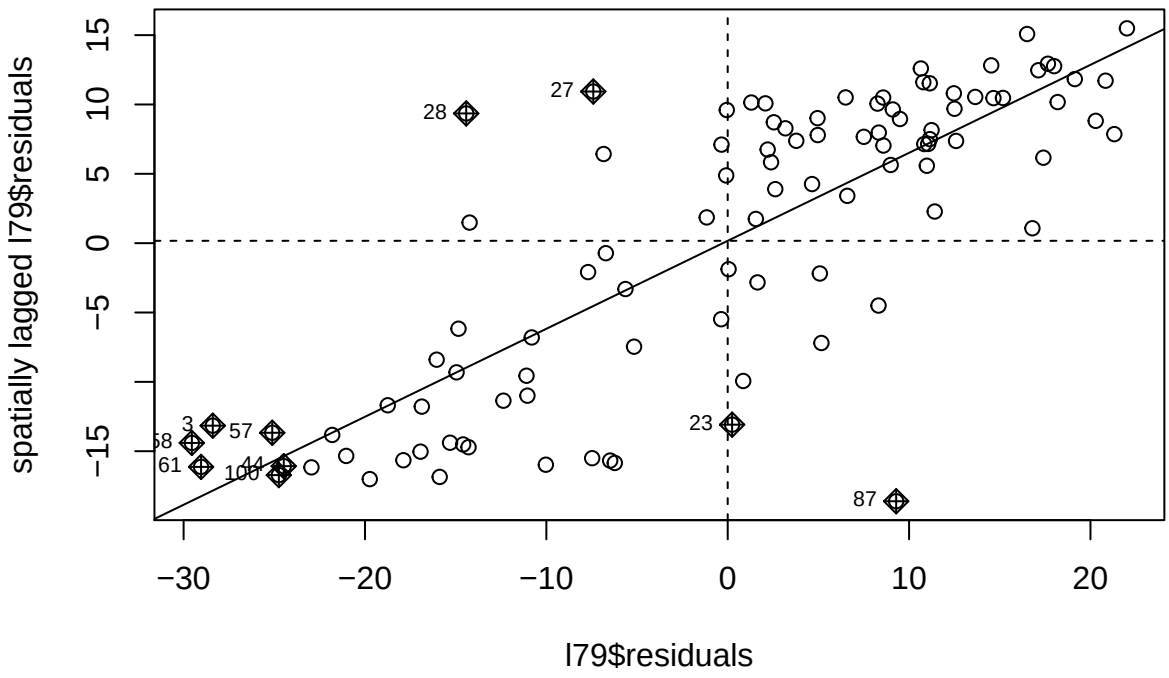

```
moran.plot(l79$residuals, listw=col.rb)
```



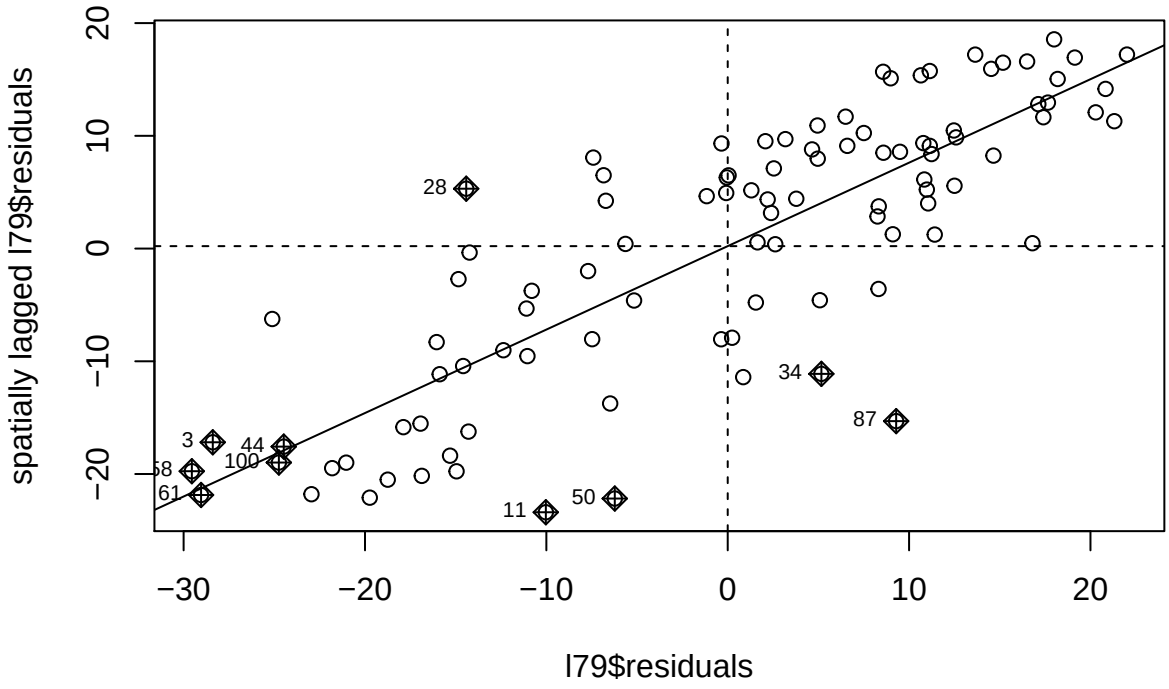

```
moran.plot(l79$residuals, listw=nc.d2)
```



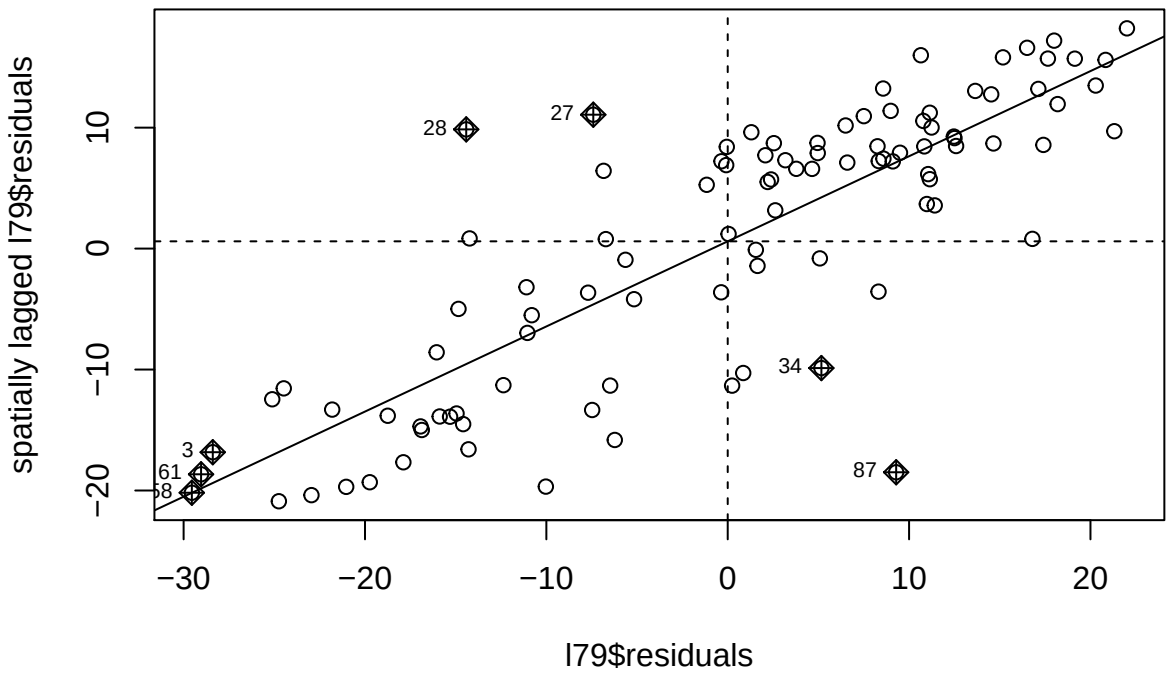

```
moran.plot(l79$residuals, listw=nc.d1)
```



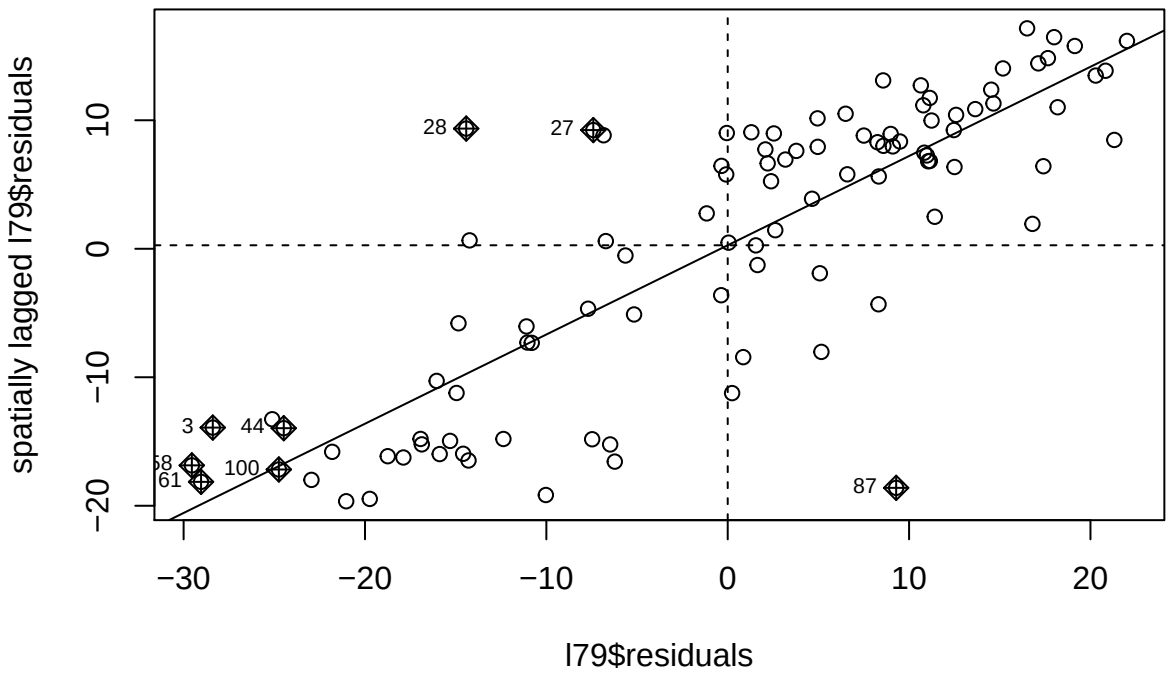

```
moran.plot(l79$residuals, listw=knn3)
```



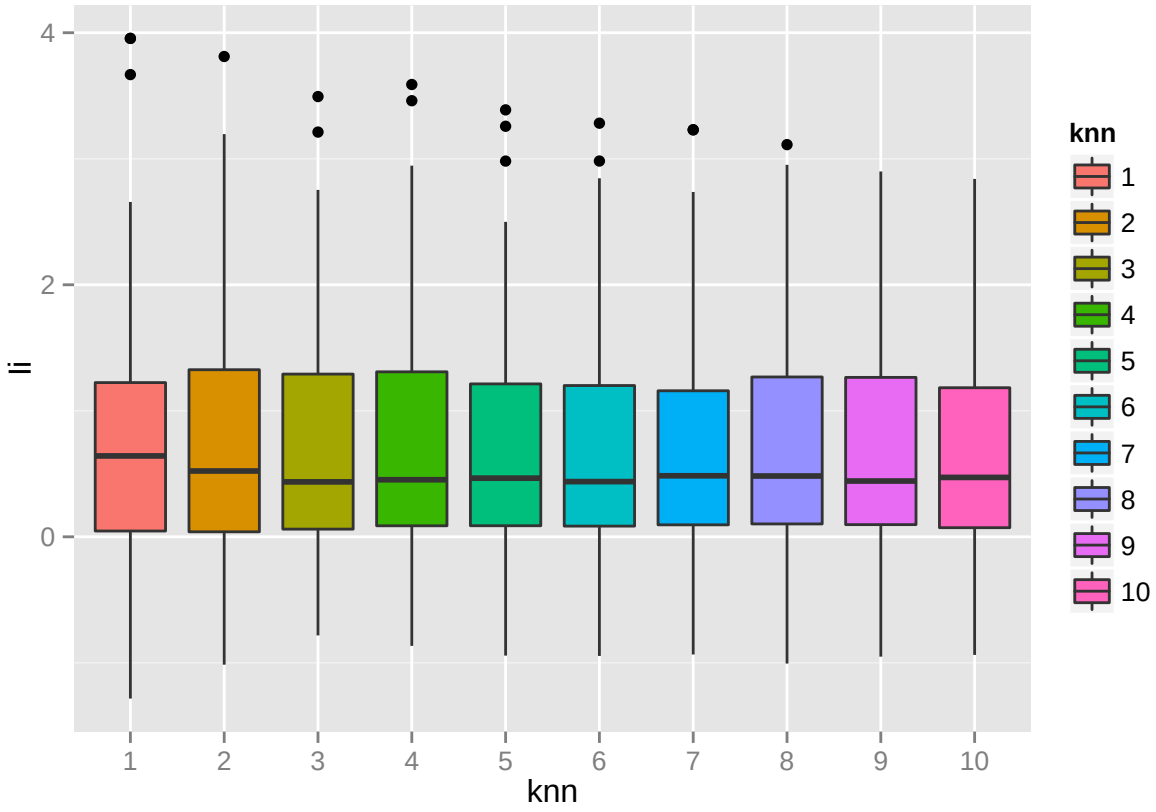

```
moran.plot(l79$residuals, listw=knn6)
```




```
moran.plot(l79$residuals, listw=knn9)
```




```
mor1 <- as.data.frame(localmoran(x=l79$residuals, listw=knn1))
mor1$knn <- 1
mor2 <- as.data.frame(localmoran(x=l79$residuals, listw=knn2))
mor2$knn <- 2
mor3 <- as.data.frame(localmoran(x=l79$residuals, listw=knn3))
mor3$knn <- 3
mor4 <- as.data.frame(localmoran(x=l79$residuals, listw=knn4))
mor4$knn <- 4
mor5 <- as.data.frame(localmoran(x=l79$residuals, listw=knn5))
mor5$knn <- 5
mor6 <- as.data.frame(localmoran(x=l79$residuals, listw=knn6))
mor6$knn <- 6
mor7 <- as.data.frame(localmoran(x=l79$residuals, listw=knn7))
mor7$knn <- 7
mor8 <- as.data.frame(localmoran(x=l79$residuals, listw=knn8))
mor8$knn <- 8
mor9 <- as.data.frame(localmoran(x=l79$residuals, listw=knn9))
mor9$knn <- 9
mor10 <- as.data.frame(localmoran(x=l79$residuals, listw=knn10))
mor10$knn <- 10

mor <- rbind(mor1, mor2, mor3, mor4, mor5,mor6,mor7,mor8,mor9,mor10)
mor$knn <- as.factor(mor$knn)
ggplot(data=mor)+geom_boxplot(aes(knn,Ii, fill=knn, group=knn))
```



what about distance from the ocean? elevation?