Dem 7263
Spatial Autocorrelation in R

#Set your working directory
setwd("/courses/dem7263cs/data")

#load libraries, you will need to install these
library(spdep)
library(maptools)

dat<-readShapePoly("satract.shp", proj4string=CRS("+proj=utm zone=14"))
saol<-readShapeLines( "sa_l.shp", proj4string=CRS("+proj=utm zone=14"))

#do a VERY simple plot of the tracts
plot(dat)

#do a slightly more complicated map of the property crime rate
spplot(dat, "pcrimert")
Which basically just identifies one area of downtown as having LOTS of crimes per captia. Let’s make a little more difficult, but more informative map

library(classInt)
library(RColorBrewer)
brks<-classIntervals(dat$pcrimert, n=5,style="quantile")
cols<-brewer.pal(5, "Blues")
plot(saol, col=grey(.1))
plot(dat, col=cols[findInterval(dat$pcrimert, brks$brks, all.inside=T)], border=0,add=T, main="Property Crime Rate /1000 People")
legend("bottomleft", legend=leglabs(round(brks$brks,1), under="<", over=">", between="-"), fill=cols, title="Rate", cex=1.25)
# Now let’s calculate Moran’s I, first we need a spatial weight matrix
# Make a rook style weight matrix
sanb <- poly2nb(dat, queen = F)
summary(sanb)

Neighbour list object:
Number of regions: 235
Number of nonzero links: 1106
Percentage nonzero weights: 2.002716
Average number of links: 4.706383
Link number distribution:

1 2 3 4 5 6 7 8 9
4 10 30 62 66 34 24 3 2
4 least connected regions:
61 82 147 205 with 1 link
2 most connected regions:
31 55 with 9 links
salw<-nb2listw(sanb, style="W")
#the style = W row-standardizes the matrix

#Calculate univariate moran's I for the property crime rate
moran.test(dat$pcrimert, listw=salw)
  Moran's I test under randomisation

data:  dat$pcrimert
weights: salw

Moran I statistic standard deviate = 2.2194, p-value = 0.01323 ↵

alternative hypothesis: greater
sample estimates:
  Moran I statistic   Expectation   Variance
   0.087742862  -0.004273504  0.001718933

#Calculate the test using Monte Carlo sampling, slightly better, this is like GeoDa
moran.mc(dat$pcrimert, listw=salw, nsim=999)
Monte-Carlo simulation of Moran's I

data:  dat$pcrimert
weights: salw
number of simulations + 1: 1000

statistic = 0.0877, observed rank = 977, p-value = 0.023 ↵
alternative hypothesis: greater

#Moran scatter plot
moran.plot(dat$pcrimert, listw=salw)
par(mfrow=c(2,1))
plot(sp.correlogram(dat$pcrimert, neighbours=sanb, order=5))
plot(sp.correlogram(dat$ppersonsp, neighbours=sanb, order=5))
The Crime rate doesn’t show much autocorrelation, and certainly low at higher order lags, while poverty shows high autocorrelation even among distant tracts.