rm(list = ls())

setwd("C:/Users/tuj53509/Dropbox/docs/Temple/Advanced Statistics for Urban Applications/Assignment2")

library(GISTools)

library(maptools)

library(spgwr) #this is a package for conducting gwr

library(spdep) # this package allows for testing of spatial autocorrelation (dependency)

library(lm.beta)

library(pastecs)

#disable scientific notation

options(scipen=999)

kentucky = readShapeSpatial("ky\_counties/ky\_counties", proj4string=CRS("+proj=lcc"))

kentucky\_census = read.csv("KY.bband.csv")

# Join census file to shapefile

kentucky@data <- data.frame(kentucky@data, kentucky\_census[match(kentucky@data[,"NAME10"],kentucky\_census[,"Name"]),])

#attach to the data frame to simplify calling vectors

attach(data.frame(kentucky))

#Step 3: Visualize and explore data

#add descriptive statistics

#add a second visualization

vals <- cbind(adopt05, college, hhpct\_0401, popden2000)

stat.desc(vals)

pdf(file = "f1.adopt05histogram.pdf")

hist(adopt05, col = "lightgreen", xlim = c(0,1), main = "Figure 1: Broadband Adoption by County", xlab = "proportion of adoption")

dev.off()

#choropleth of adopt05 (plot1)

par(mar = c(0,0,0,0))

pdf(file="f2.adopt05choropleth.pdf")

adopt05.shades <- auto.shading(adopt05, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, adopt05, shading = adopt05.shades)

title("Figure 2: Kentucky Broadband Adoption")

choro.legend(3751596,4833529, adopt05.shades, title = "proportion of adoption")

map.scale(5021404, 3205804, miles2ft(200), "Miles", 5, 40)

north.arrow(5898350, 4350232, miles2ft(10))

dev.off()

#Step 4: Linear Model and Diagnostics

#model

lm01 <- lm(adopt05 ~ college + hhpct\_0401 + popden2000)

#add standardized betas

lm01.std <- lm.beta(lm01)

summary(lm01.std)

summary(lm01$residuals)

#add residuals to the kentucky data frame

kentucky$lm01resid<-residuals(lm01)

#plot residuals (plot2)

pdf(file = "f3.lm01residchoropleth.pdf")

lm01residshades = auto.shading(kentucky$lm01resid, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, kentucky$lm01resid, shading = lm01residshades)

title("Figure 3: Residuals from linear model lm01")

choro.legend(3751596,4833529, lm01residshades, title = "residuals")

dev.off()

#neighbor plot with distances

kentucky\_neighbors <- poly2nb(kentucky)

pdf(file = "f4.weightedneighborplot.pdf")

plot(kentucky, main = "Figure 4: Weighted Neighbor plot of Kentucky Broadband Adoption ")

plot(kentucky\_neighbors, coordinates(kentucky), add=T, col='blue')

dev.off()

#Lagged Means

KY\_Neigh\_lw <- nb2listw(kentucky\_neighbors)

bband <- adopt05

bband\_lagged\_mean <- lag.listw(KY\_Neigh\_lw, bband)

par(mar = c(2,4,4,2))

pdf(file="f5.lagged\_means\_choropleth.pdf")

laggedmeanshades = auto.shading(bband\_lagged\_mean, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, bband\_lagged\_mean, shading =laggedmeanshades)

title("Figure 5: Lagged Means plot of Kentucky Broadband Adoption")

choro.legend(3751596,4833529, laggedmeanshades, title = "Lagged Means")

dev.off()

#Moran's I - Claude Schrader

#I'm using the short version of adopt05 because I've already attached to the data

#frame. If you haven't done this, you probably need to use the full kentucky$adopt05 notation

#this calculates the maximum and minimum I values

moran.range <- function(lw) {

wmat <- listw2mat(lw)

return(range(eigen((wmat + t(wmat))/2)$values))

}

moran.range(KY\_Neigh\_lw)

#approximate test statistic using normal distribution

moran.test(adopt05, KY\_Neigh\_lw)

#calculate the test statistic using 10,000 random trials

moran.mc(adopt05, KY\_Neigh\_lw, 10000)

#this saves the plot to your working directory rather than displaying it inline

pdf(file = "f6.moransplot.pdf")

moran.plot(adopt05, KY\_Neigh\_lw)

title("Figure 6: Moran's Plot")

#this closes the file handle

dev.off()

#Step 5: GWR - Claude Schrader

kentucky.bw <- gwr.sel(adopt05 ~ college + hhpct\_0401 + popden2000, data = kentucky, gweight = gwr.Gauss)

kentucky.gwr <- gwr(adopt05 ~ college + hhpct\_0401 + popden2000, data = kentucky, bandwidth = kentucky.bw, gweight = gwr.Gauss, hatmatrix = T)

summary(kentucky.gwr)

names(kentucky.gwr)

names(kentucky.gwr$SDF)

gwr.df <- slot(kentucky.gwr$SDF, "data")

#Step 6: analyze the result of the GWR

print(kentucky.gwr)

anova(kentucky.gwr)

#Step 7: Plot Results

#spatial distribution of parameter estimates (for each X) (plot3)

#p-values (you can recode these as below and above .05)(plot5)

#calculate the t-score

gwr.df$college\_tval <- gwr.df$college / gwr.df$college\_se

#calculate the p-value

gwr.df$college\_pval <- pt(gwr.df$college\_tval, 119, lower.tail = FALSE)

gwr.df$popden2000\_tval <- gwr.df$popden2000 / gwr.df$popden2000\_se

gwr.df$popden2000\_pval <- pt(gwr.df$popden2000\_tval, 119, lower.tail = FALSE)

gwr.df$hhpct\_0401\_tval <- gwr.df$hhpct\_0401 / gwr.df$hhpct\_0401\_se

gwr.df$hhpct\_0401\_pval <- pt(gwr.df$hhpct\_0401\_tval, 119, lower.tail = FALSE)

pdf(file = "f7.collegeparamchoropleth.pdf")

collegeparamshades = auto.shading(gwr.df$college, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, gwr.df$college, shading=collegeparamshades)

title("Figure 7: GWR Parameter Estimate for College Graduation Rate")

choro.legend(3751596,4833529, collegeparamshades, title = "graduation rate")

dev.off()

pdf(file = "f8.collegepvalchoro.pdf")

collegepvalshades <- shading(c(0.005, 0.01, 0.05), cols = rev(brewer.pal(4, "Greens")))

#collegepvalshades = auto.shading(gwr.df$college\_pval, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, gwr.df$college\_pval, shading = collegepvalshades)

title("Figure 8: P-Value for College Graduation Rate")

choro.legend(3751596,4833529, collegepvalshades, title = "p-value")

dev.off()

pdf(file = "f9.popdenchoropleth.pdf")

popdenparamshades = auto.shading(gwr.df$popden2000, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, gwr.df$popden2000, shading=popdenparamshades)

title("Figure 9: GWR Parameter Estimate for Population Density in 2000")

choro.legend(3751596,4833529, popdenparamshades, title = "Population Density")

dev.off()

pdf(file = "f10.popdenpvalchoro.pdf")

popdenpvalshades = shading(c(0.005, 0.01, 0.05), cols=rev(brewer.pal(5,"Greens")))

#popdenpvalshades = auto.shading(gwr.df$popden2000\_pval, cols=rev(brewer.pal(5,"Greens")))

choropleth(kentucky, gwr.df$popden2000\_pval, shading = popdenpvalshades)

title("Figure 10: P-Value for Population Density in 2000")

choro.legend(3751596,4833529, popdenpvalshades, title = "p-value")

dev.off()

pdf(file = "f11.hhpctchoropleth.pdf")

hhpctshades = auto.shading(gwr.df$hhpct\_0401, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, gwr.df$hhpct\_0401, shading=hhpctshades)

title("Figure 11: GWR Parameter Estimate for Broadband Availability Rate")

choro.legend(3751596,4833529, hhpctshades, title = "availability")

dev.off()

pdf(file = "f12.hhpctpvalchoro.pdf")

hhpctpvalshades = shading(c(0.005, 0.01, 0.05), cols=rev(brewer.pal(5,"Greens")))

#hhpctpvalshades = auto.shading(gwr.df$hhpct\_0401\_pval, cols=rev(brewer.pal(5,"Greens")))

choropleth(kentucky, gwr.df$hhpct\_0401\_pval, shading = hhpctpvalshades)

title("Figure 12: P-Value for Broadband Availability Rate")

choro.legend(3751596,4833529, hhpctpvalshades, title = "p-value")

dev.off()

#spatial distribution of local r^2 (plot4)

pdf(file = "f13.localr2.pdf")

localr2shades = auto.shading(gwr.df$localR2, cols=brewer.pal(5,"Greens"))

choropleth(kentucky, gwr.df$localR2, shading=localr2shades)

title("Figure 13: GWR Local R² values")

choro.legend(3751596,4833529, localr2shades, title = "Local R²")

dev.off()

#don't forget to detach the data frame

detach(data.frame(kentucky))

#reenable scientific notation

options(scipen=0)