

Spatial Voting Patterns of 2016 Election in Minnesota

Nick Brown, Landon Thompson, Ali Lacey, Kyle Lifferth

December 6, 2018

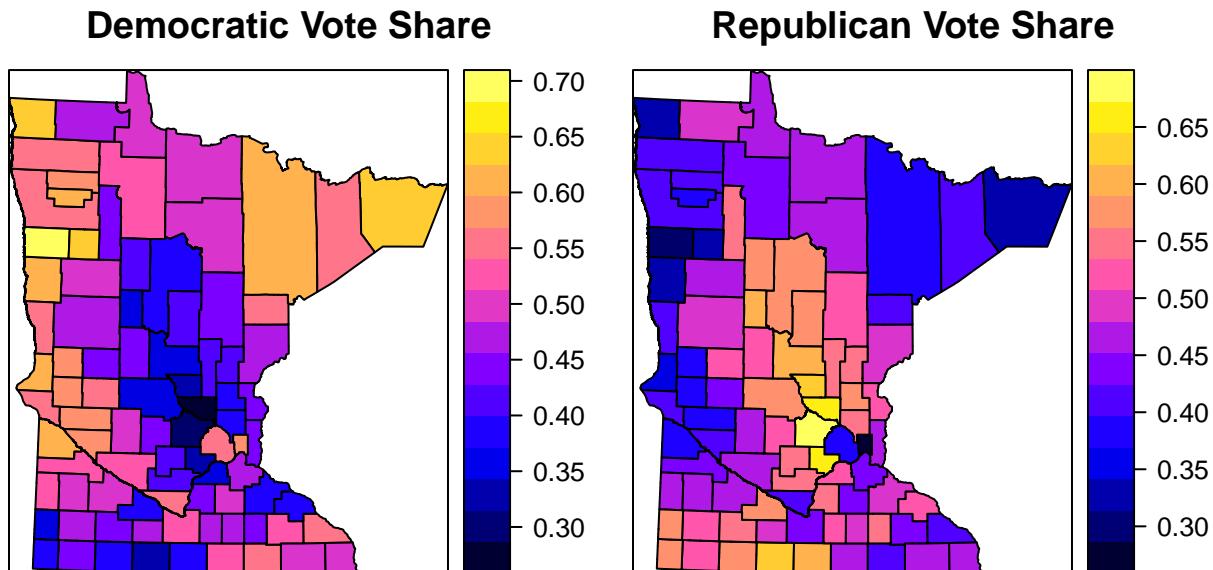
Introduction

Our data set contains voter information and voting patterns by county. When analyzing these types of data, there are generally two common goals. The first is to attempt to project the outcome of future elections. The second is to use this data to plan electoral strategy. Our analysis will focus primarily on planning electoral strategy for a state-wide race in the state of Minnesota, whether it be a governor's race or a United States Senate race.

Exploratory Analysis

To try to plan an electoral strategy, it is vital to know understand and identify any potential patterns in the data. We begin by analyzing the Vote Share Percentage by Political party.

```
Partyplot1 <- spplot(county.data, "percDem", main='Democratic Vote Share')
Partyplot2 <- spplot(county.data, "percRep", main='Republican Vote Share')
Partyplots <- list(Partyplot1,Partyplot2)
do.call(grid.arrange,c(Partyplots,nrow=1))
```

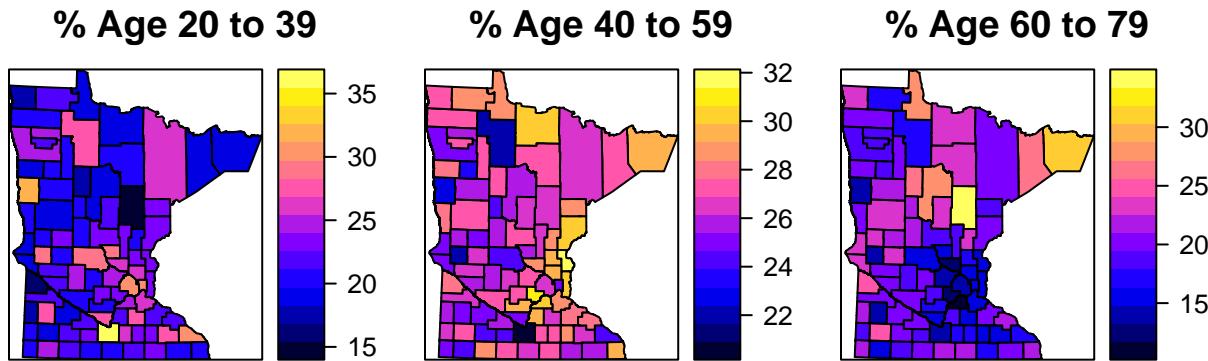


Brighter colors on the plots indicate higher than 50 percent of the vote went to that party and darker colors

indicate that less than 50 percent of the vote. These plots provided some expected and some surprising results compared to our intuition about Minnesota politics. For example, our intuition led us to believe that Republicans would have done well in the suburbs of Minneapolis. In general, suburban areas lean more conservative politically. The unexpected results was how well Democrats did in the northwestern part of the state. Our intuition led us to believe that the western parts of Minnesota would look more like South Dakota and North Dakota politically, but in the 2016 election, this was not the case.

The next focus of our analysis was the distribution of age across the state.

```
Ageplot1 <- spplot(county.data, "Age60to79", main='% Age 60 to 79')
Ageplot2 <- spplot(county.data, "Age40to59", main='% Age 40 to 59')
Ageplot3 <- spplot(county.data, "Age20to39", main='% Age 20 to 39')
Ageplots <- list(Ageplot3, Ageplot2, Ageplot1)
do.call(grid.arrange, c(Ageplots, nrow=1))
```

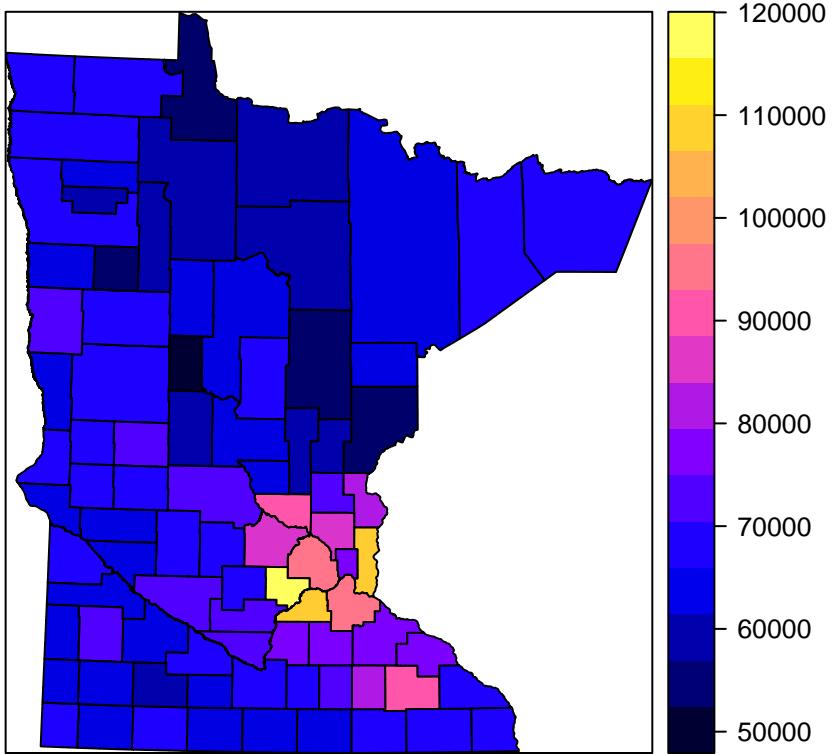


These plots matched our intuition quite closely. Younger people were concentrated in the Minneapolis metro area, as well as in the major colleges throughout the state. Moorhead State and Concordia in the northwest corner, Southwest Minnesota State in the southwest corner, and Mankato State in south central Minnesota. Older people were more concentrated in rural areas, while the highest concentrations of people 40-59 are found in the suburbs.

Next we explored how the mean income was distributed among the counties of Minnesota, measured in dollars.

```
spplot(county.data, "MeanIncome", main='Mean Income by County')
```

Mean Income by County

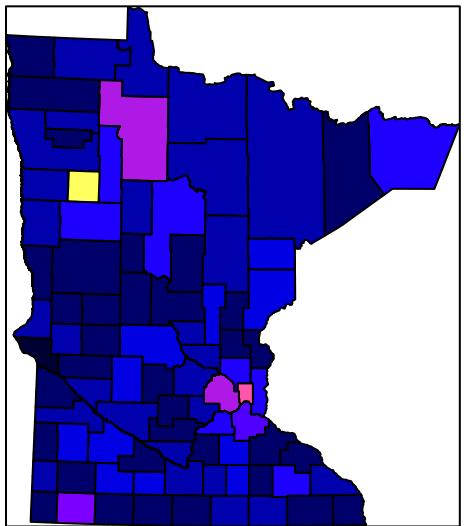


Again, the distribution of income is what we expected. A majority of the wealth is concentrated in the Minneapolis metro area and Rochester area. This is because of the professional and white collar job opportunities that generally pay higher than the blue collar jobs and agricultural jobs that dominate the rest of the state.

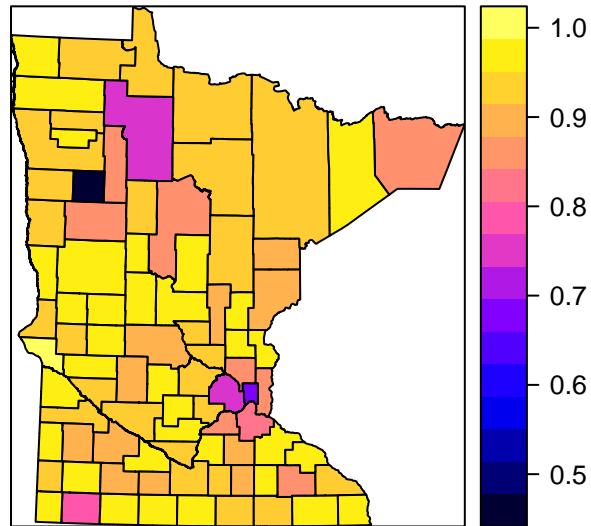
Our final exploratory category, is to see how Minnesotans are distributed according to race. We have data for various minority groups, but we wanted to compare the percentage of minority populations versus percentage of white population of Minnesota.

```
county.data@data$perc_minority <- 1-county.data@data$perc_white
Raceplot1 <- spplot(county.data,"perc_minority", main='Percent of Population - Minority')
Raceplot2 <- spplot(county.data,"perc_white", main='Percent of Population - White')
Raceplots <- list(Raceplot1,Raceplot2)
do.call(grid.arrange,c(Raceplots,nrow=1))
```

Percent of Population – Minority



Percent of Population – White



The distribution of the population in Minnesota by race was also what we expected. A majority of the counties are primarily white, with just a few exceptions, mainly being the Minneapolis and Saint Paul areas, as well as the main Native American reservation in the northwest corner of the state.

Autocorrelation Analysis By Party Vote Share

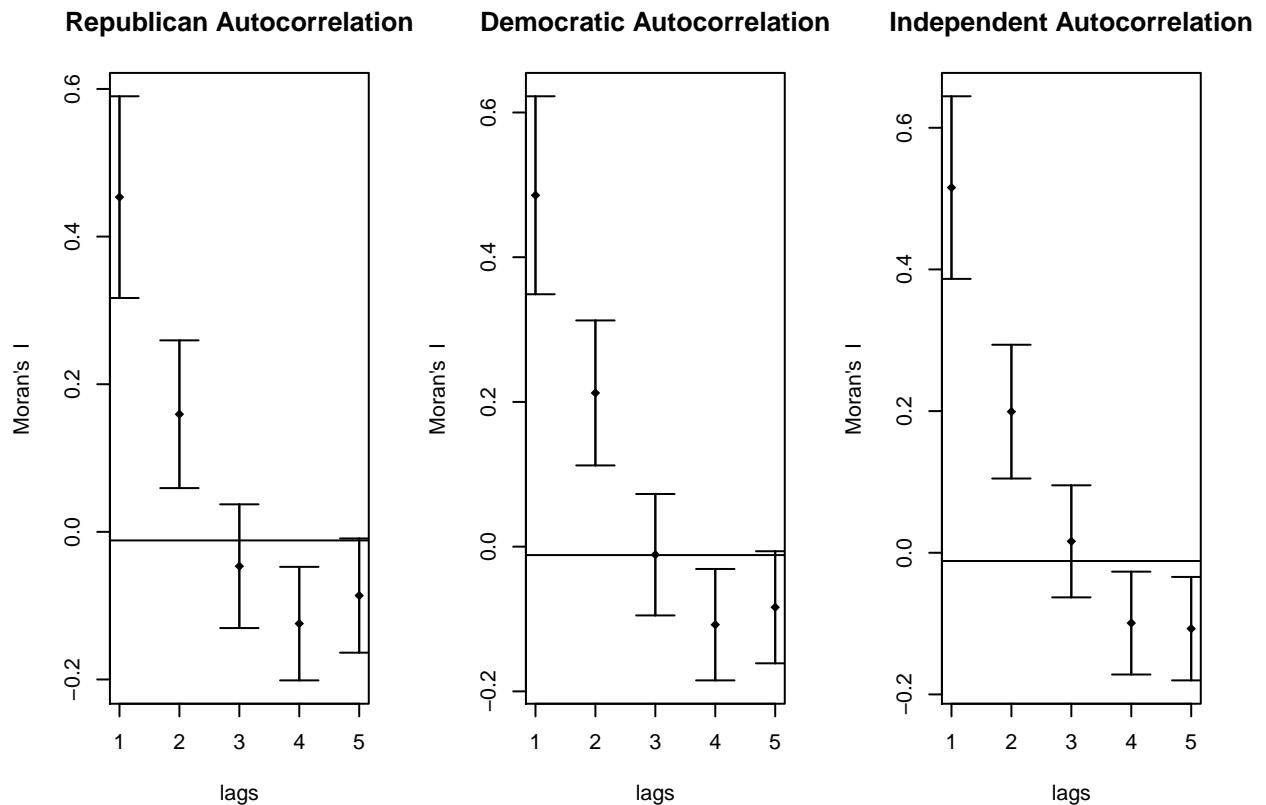
Now that we have explored the different demographic patterns in the state, we decided to look at how Voting patterns by party were spatially autocorrelated if at all. We begin with looking at Global Moran's I test for spatial autocorrelation.

```
nlist.queen <- poly2nb(county.data, row.names=as.character(county.data@data$SP_ID), queen=T)
W <- nb2listw(nlist.queen, style="W")
percRep <- county.data@data$percRep
RepMoran <- as.data.frame(localmoran(percRep,W))
colnames(RepMoran) <- c("I","EI","VarI","Zi","Pval")
Rep <- county.data
Rep@data <- RepMoran
Rep@data$percRep <- county.data@data$percRep
Rep@data$PvalFlag <- ifelse(Rep@data$Pval < .05, 1,0)

percDem <- county.data@data$percDem
DemMoran <- as.data.frame(localmoran(percDem,W))
colnames(DemMoran) <- c("I","EI","VarI","Zi","Pval")
Dem <- county.data
Dem@data <- DemMoran
Dem@data$percDem <- county.data@data$percDem
Dem@data$PvalFlag <- ifelse(Dem@data$Pval<.05,1,0)

percInd <- county.data@data$percInd
IndMoran <- as.data.frame(localmoran(percInd,W))
colnames(IndMoran) <- c("I","EI","VarI","Zi","Pval")
Ind <- county.data
Ind@data <- IndMoran
Ind@data$percInd <- county.data@data$percInd
Ind@data$PvalFlag <- ifelse(Ind@data$Pval <0.05,1,0)

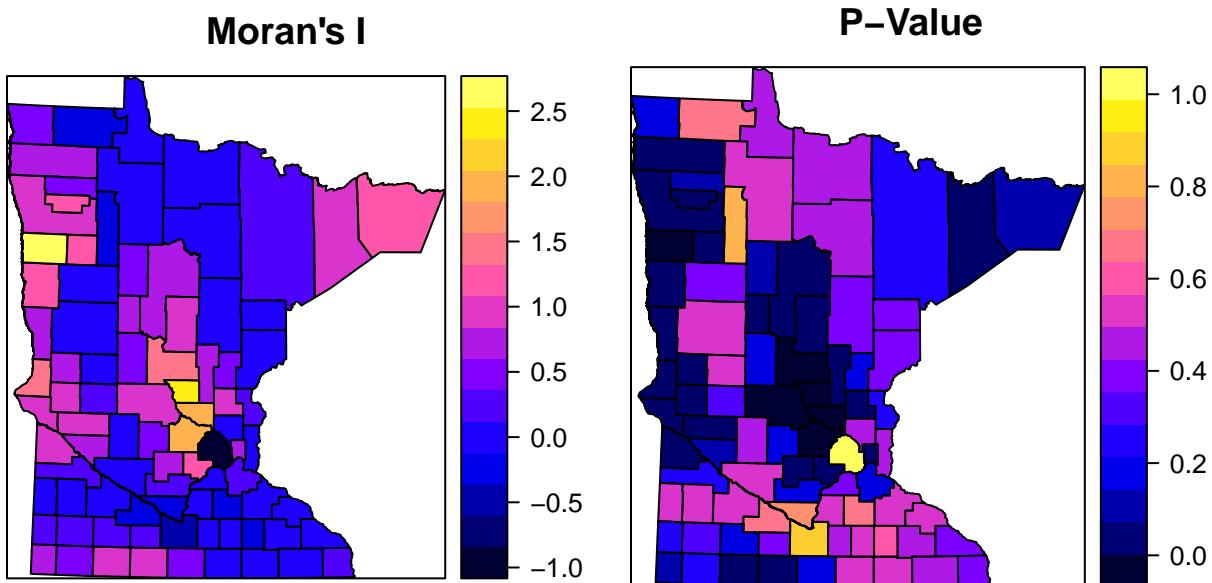
Rep.q <- sp.correlogram(nlist.queen,Rep@data$percRep,order=5,method="I",style="W",
                        randomisation = T)
Dem.q <- sp.correlogram(nlist.queen,Dem@data$percDem,order=5,method="I",style="W",
                        randomisation=T)
Ind.q <- sp.correlogram(nlist.queen,Ind@data$percInd,order=5,method="I",style="W",
                        randomisation=T)
par(mfrow=c(1,3))
plot(Rep.q,main="Republican Autocorrelation")
plot(Dem.q, main="Democratic Autocorrelation")
plot(Ind.q, main="Independent Autocorrelation")
```



For Democratic and Independent vote shares, the Global Moran's I test shows positive spatial autocorrelation at one and two lag distances away. Here a lag distance is one or two counties away. This would lead us to believe that there is some kind of clustering of counties that are all spatially autocorrelated. Meanwhile, Republican vote share only has positive spatial autocorrelation at lag distance one, or one county away. This leads us to think that there might be a number of pairs of counties that are positively spatially autocorrelated, but there might not be large groups.

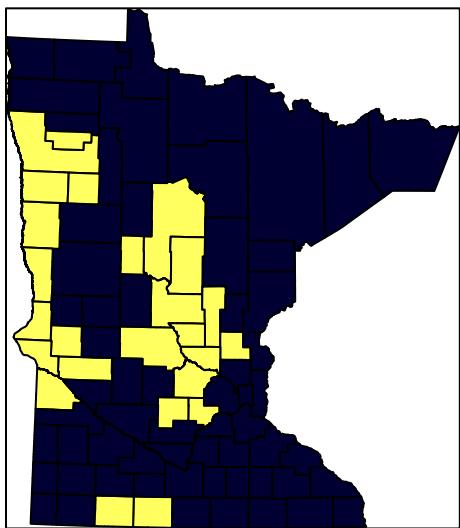
Now considering local measures of spatial autocorrelation, we conduct hypothesis testing using Local Moran's I, using the queens case to determine neighbors, to check for spatial autocorrelation by county for percent vote share by party, first for the Republican Party.

```
Repplot3 <- spplot(Rep, "I", main="Moran's I")
Repplot4 <- spplot(Rep, "Pval", main="P-Value")
Replots2 <- list(Repplot3,Repplot4)
do.call(grid.arrange,c(Replots2,nrow=1))
```

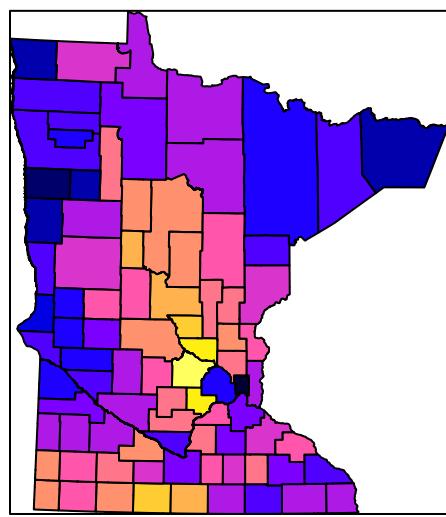


```
Repplot1 <- spplot(Rep, "PvalFlag",main="Republican Vote Autocorrelation")
Repplot2 <- spplot(Rep, "percRep",main="Republican Vote Share")
Replots <- list(Repplot1,Repplot2)
do.call(grid.arrange,c(Replots,nrow=1))
```

Republican Vote Autocorrelation



Republican Vote Share



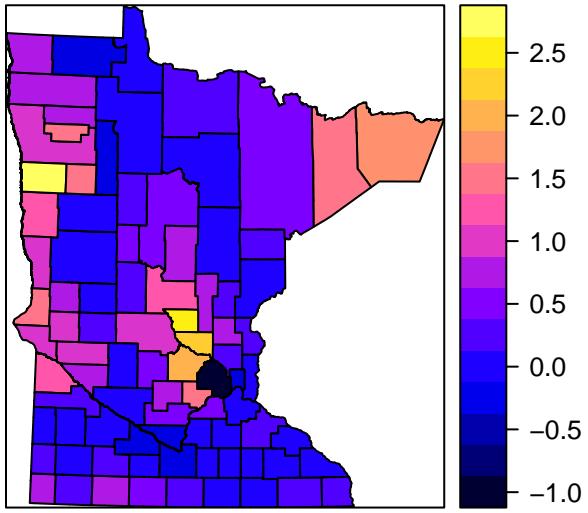
The yellow counties mean that there is significant positive spatial autocorrelation. Notice that counties that exhibit positive spatial autocorrelation have one of two properties. Counties that had a high percentage of Republican votes and counties that had a low percentage of Republican votes. Another interesting observation is that the Moran's I statistic for Hennepin county is close to -1. This is likely close to being negatively spatially autocorrelated since it is so drastically different with respect to Republican vote share.

We also note that despite what the global measure of autocorrelation showed, there are indeed groupings of counties that are spatially autocorrelated with respect to Republican Vote share. But this might suggest that even though each individual county highlighted in yellow is spatially autocorrelated, the relationship is only with respect to the county next to it and not necessarily based on two counties away.

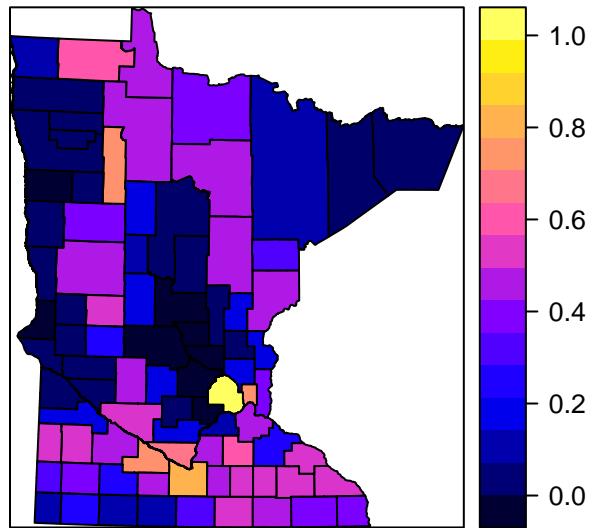
Now let's move on to considering the Democratic vote share under the same test and procedure.

```
Demplot3 <- spplot(Dem, "I", main="Moran's I")
Demplot4 <- spplot(Dem, "Pval", main="P-Value")
Demplots2 <- list(Demplot3,Demplot4)
do.call(grid.arrange,c(Demplots2,nrow=1))
```

Moran's I

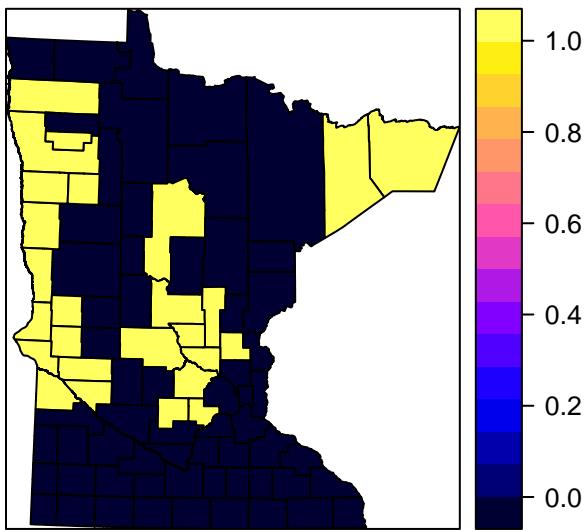


P-Value

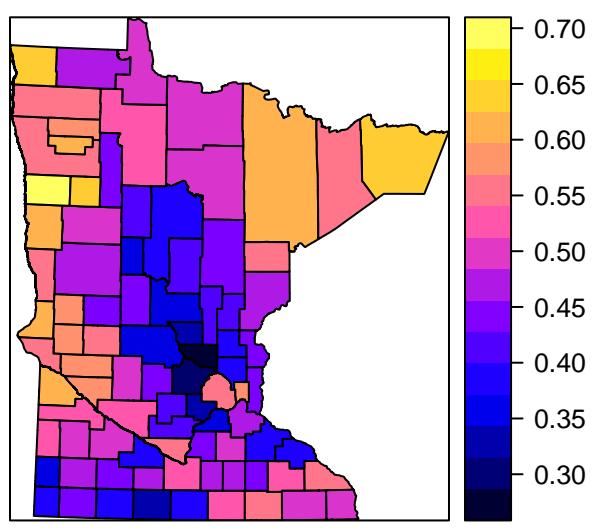


```
Demplot1 <- spplot(Dem, "PvalFlag", main="Democratic Vote Autocorrelation")
Demplot2 <- spplot(Dem, "percDem", main="Democratic Vote Share")
Demplots <- list(Demplot1,Demplot2)
do.call(grid.arrange,c(Demplots,nrow=1))
```

Democratic Vote Autocorrelation



Democratic Vote Share



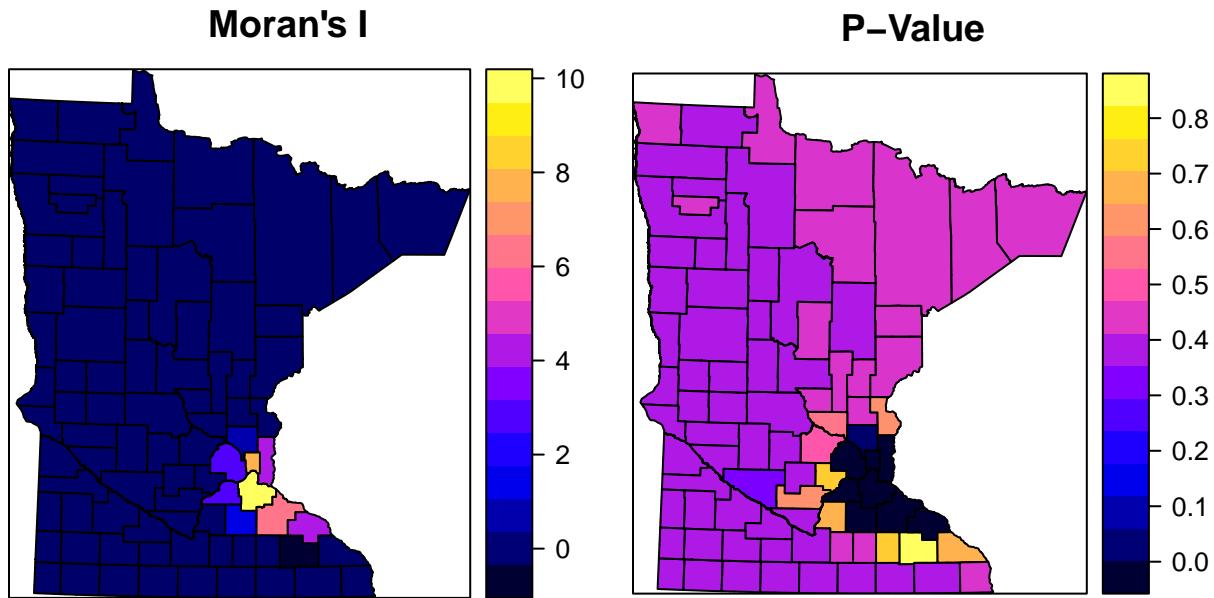
Here we see very similar trends to the Republican Vote share, both in terms of spatial autocorrelation and clusters of counties. The clusters of counties are more expected here though after the test for global autocorrelation suggested that there was positive spatial autocorrelation at lag distance two.

Finally we considered the Independent and Write-in vote shares under the same test and procedure.

```

Indplot3 <- spplot(Ind, "I", main="Moran's I")
Indplot4 <- spplot(Ind, "Pval", main="P-Value")
Indplots2 <- list(Indplot3,Indplot4)
do.call(grid.arrange,c(Indplots2,nrow=1))

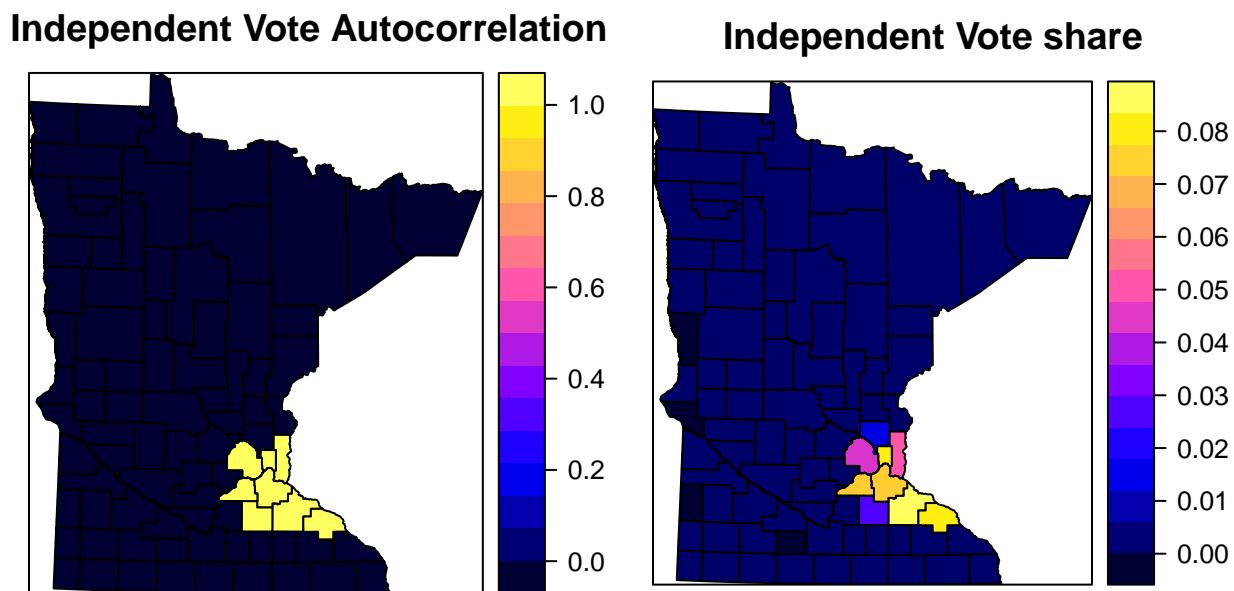
```



```

Indplot1 <- spplot(Ind,"PvalFlag",main="Independent Vote Autocorrelation")
Indplot2 <- spplot(Ind,"percInd",main="Independent Vote share")
Indplots <- list(Indplot1,Indplot2)
do.call(grid.arrange,c(Indplots,nrow=1))

```



After seeing the pattern that had emerged when testing for spatial autocorrelation for Democratic Vote

Share and Republican Vote Share, the results for Independents was slightly less surprising. What is notable however, is that its only in the metro area of Minneapolis that there was any significant Independent Vote share. This could be because of the diversity of the populations that live there. Meanwhile, most of the other states had very few votes for Independents.

Conclusion

In conclusion, analyzing spatial autocorrelation based on party vote share, provides us with interesting insights as to how to campaign for a statewide race in Minnesota. Knowing where you do well and where counties are spatially autocorrelated would inform strategy about how and where to campaign. It would be very beneficial to campaign in areas where historically, your party has done well, and where there is spatial autocorrelation. Many times campaigns become driven by momentum. Campaigning in suburbs for Republicans and on the western edge of Minnesota for Democrats would provide opportunities to either build momentum for your campaign or change momentum if you are doing poorly.

Another area to research for campaigns would be to use the demographic information to try to predict vote share by party. This would give you an idea of what issues to hit in particular counties to maximize the effectiveness of your campaign.