Mapping with R-Assignment 4

Code

Raghu Sanugommula and Subhash Pemmaraju

6th March 2019

Introduction: Initial Summary

The data deals with public schools in Pennsylvalnia. The questions dealt with this analysis are those of race, income levels (school lunches), level of schools and school attendance. Initial setup of the data involves loading the counties and district data, creating new variables for % white and % nonwhite and to replace all default values for population in a particular grade KG, PKG, G01-G13 from -2 to 0[1].

Question 1:

What is the distribution of White vs Non-White population across the state?

Hide

```
#1 - White vs Non-White %
nonwhiteperc = leaflet() %>% addTiles() %>%
  addCircleMarkers(data = easternschools, radius = ~nonwhite*10, color = "red", weight = 0.75, opacity = 1, fill = TRUE, fil
lColor = "red", fillOpacity = 0.2, label = ~city) %>%
  addLegend("topright", colors = c("red"), labels = c("% Non-White"), title = "Races", opacity = 1) %>%
  addRectangles(
    lng1 = -74.7,
    lat1 = 39.7,
    lng2 = -75.7,
    lat2 = 40.3,
    fillColor = "transparent",
    weight = 2.5,
    color = 'green'
  )


whiteperc = leaflet() %>% addTiles() %>%
  addCircleMarkers(data = easternschools, radius = ~white*10, color = "green", weight = 0.75, opacity = 1, fill = TRUE, fill
Color = "green", fillOpacity = 0.2, label = ~city) %>%
  addLegend("topright", colors = c("green"), labels = c("% White"), title = "Races", opacity = 1)


latticeView(nonwhiteperc,whiteperc)
```

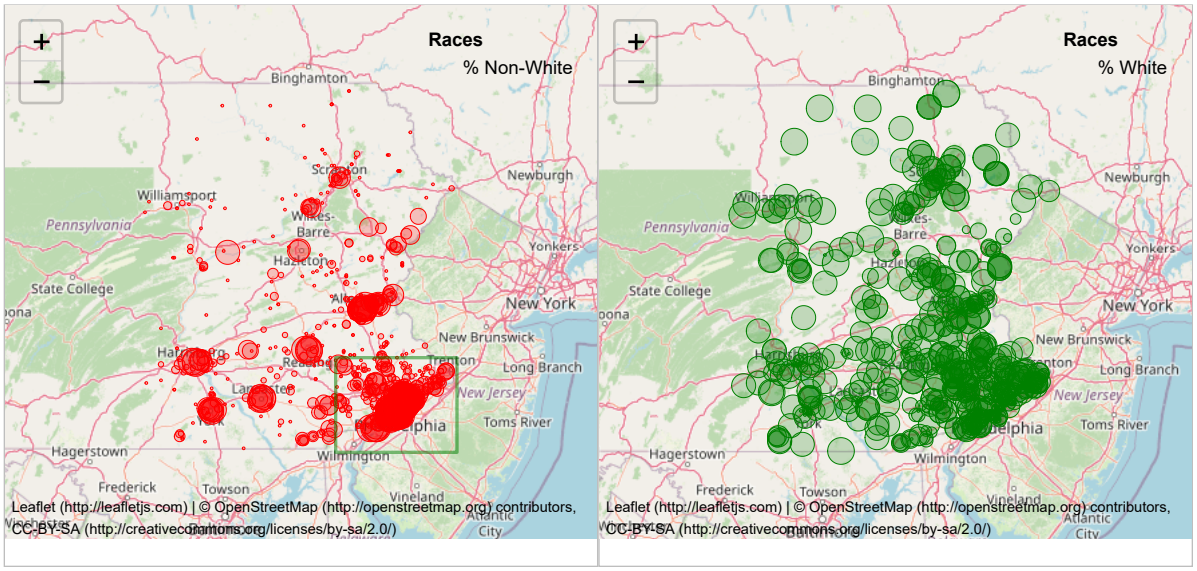


Figure 1: Distribution of White vs Non-White population across the geography

Leaflet is used for the above plot[2]. The idea behind developing the above plot is to identify where the white and non-white school population is distributed acros the geography. Although both plots are essentially of the same data points with one highlighting % nonwhites and the other identifying % whites, the idea behind showing both is to indicate the stark contrast in distribution. Non-white populations cluster together with a particularly large cluster around Philadelphia as highlighted. White population on the other hand is dispersed widely across the geography.

Question 2:

What is the distribution of levels of schooling across the state?

Hide

```
#2 - School Level
elementary = easternschools %>% filter(LEVEL==1)
middle = easternschools %>% filter(LEVEL==2)
high = easternschools %>% filter(LEVEL==3)
combined = easternschools %>% filter(LEVEL==4)

schoollevel = leaflet(counties) %>% addTiles() %>%
  setView(lat=40.2, lng=-75.5,zoom=8.5)%>%
  addPolygons(weight=1,color="blue",fillOpacity = 0.05) %>%
  addCircleMarkers(data = elementary, radius = 4, color = "magenta", weight = 0.75, opacity = 1, fill = TRUE, fillColor = "m
agenta", fillOpacity = 1, label = ~city) %>%
  addCircleMarkers(data = middle, radius = 4, color = "cyan", weight = 0.75, opacity = 1, fill = TRUE, fillColor = "cyan", f
illOpacity = 1, label = ~city) %>%
  addCircleMarkers(data = high, radius = 4, color = "green", weight = 0.75, opacity = 1, fill = TRUE, fillColor = "green", f
illOpacity = 1, label = ~city) %>%
  addLegend("topright", colors = c("magenta", "cyan", "green"), labels = c("Elementary School", "Middle School", "High Schoo
l"), title = "School Level", opacity = 1)

schoollevel
```

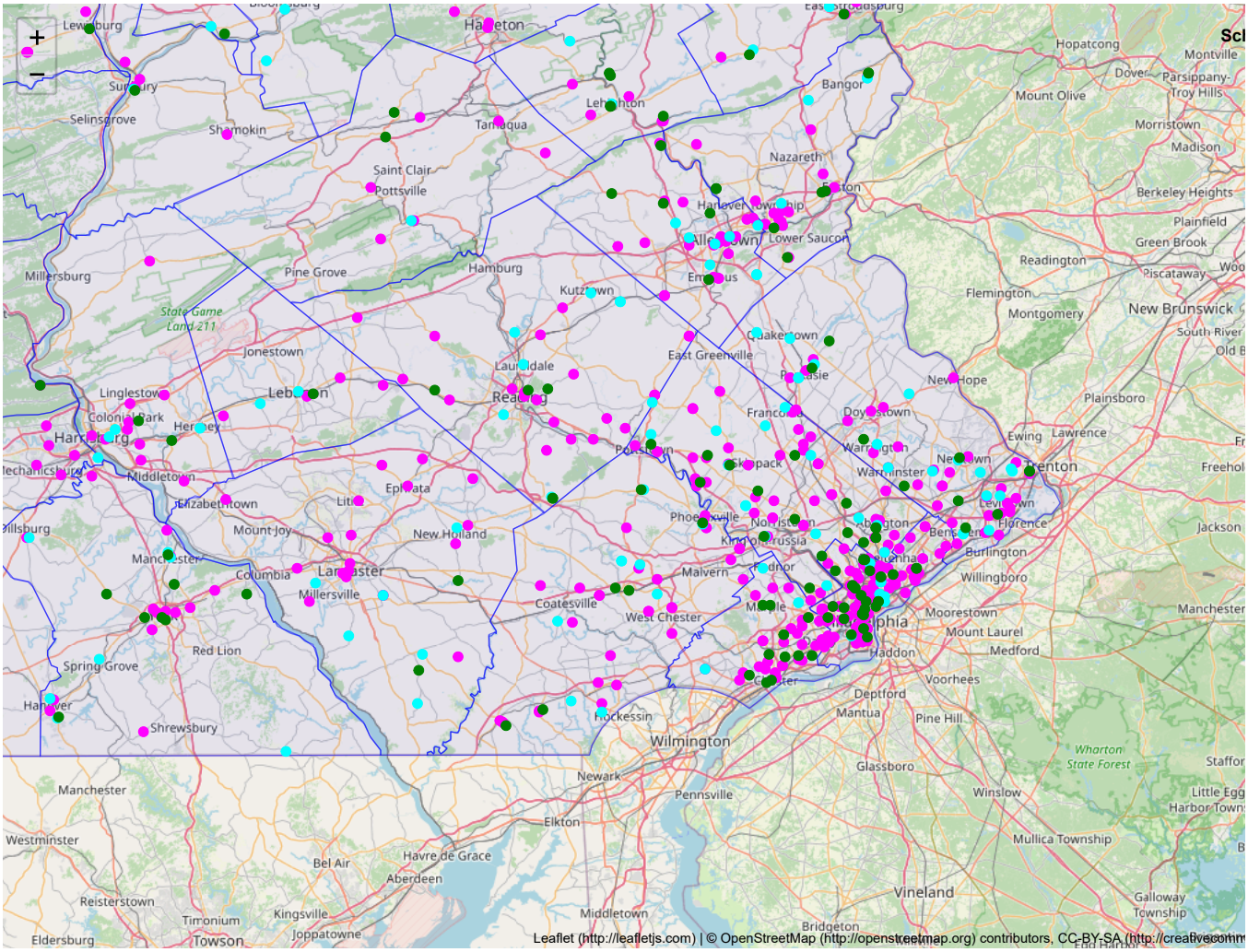


Figure 2: Distribution of levels of schools across the state

A plot of the school level distribution shows that there is heavy concentration of all levels of schooling around urban centers like phildelphia. The plot has been zoomed in to focus on a particular area using setview[3]. While the rest of the state has plenty of elementary schools and to a certain extent middle schools, high schools seem to be far and dispersed in non-urban centers, which would imply that children would have to travel a greart distance to attend high school.

Question 3:

Hide

```
#3 - school lunches
easternschools$PctFreeLunch<-as.numeric(as.character(easternschools$PctFreeLunch))
easternschools$freelunch <- ifelse(easternschools$PctFreeLunch >= 30,1,0)
freelunchesnonwhite = easternschools %>% filter(nonwhite>0.5, freelunch==1)
freeluncheswhite = easternschools %>% filter(white>0.5, freelunch==1)
schoolluncheswhite1 = leaflet(districts) %>% addTiles() %>%
#  setView(lat=40.2, lng=-75.5,zoom=8.5)%>%
#  addPolygons(weight=1,color="blue",fillOpacity = 0.05) %>%
  addCircleMarkers(data = freeluncheswhite, radius = 4, color = "red", weight = 0.75, opacity = 1, fill = TRUE, fillColor =
"red", fillOpacity = 0.2, label = ~city) %>%
  addLegend("topright", colors = c("red"), labels = c("Free Lunch %>30% in mostly White schools"), title = "Free Lunches", o
pacity = 1)

schoollunchesnonwhite1 = leaflet(districts) %>% addTiles() %>%
#  setView(lat=40.2, lng=-75.5,zoom=8.5)%>%
#  addPolygons(weight=1,color="blue",fillOpacity = 0.05) %>%
  addCircleMarkers(data = freelunchesnonwhite, radius = 4, color = "green", weight = 0.75, opacity = 1, fill = TRUE, fillCol
or = "green", fillOpacity = 0.2, label = ~city) %>%
  addLegend("topright", colors = c("green"), labels = c("Free Lunch %>30% in mostly nonWhite schools"), title = "Free Lunche
s", opacity = 1)
latticeView(schoolluncheswhite1,schoollunchesnonwhite1)
```

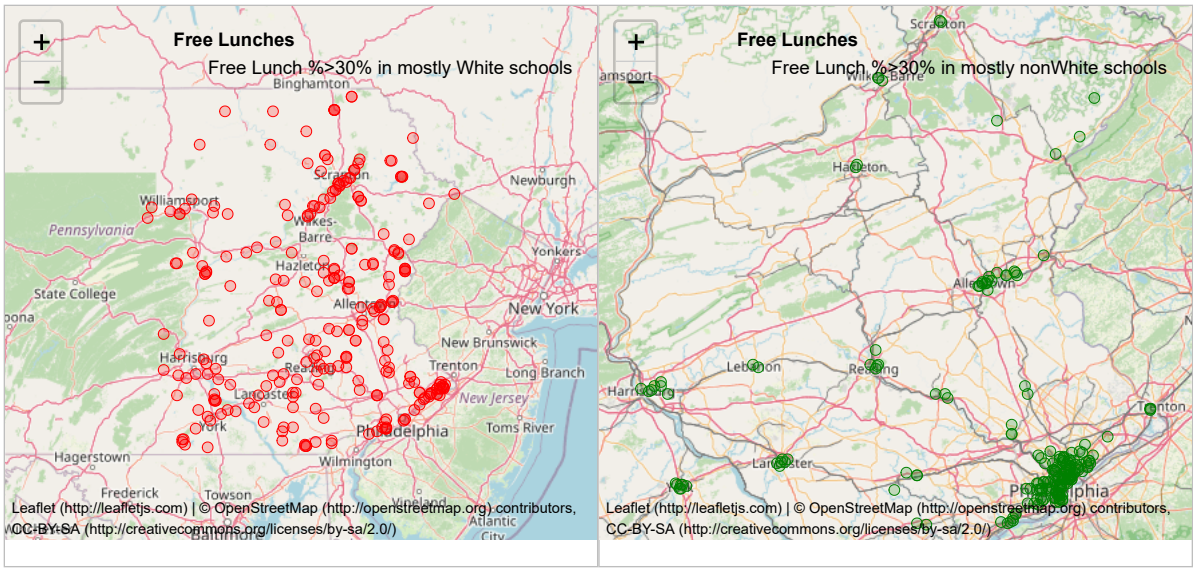


Figure 3: Distribution of free lunch recepients by race

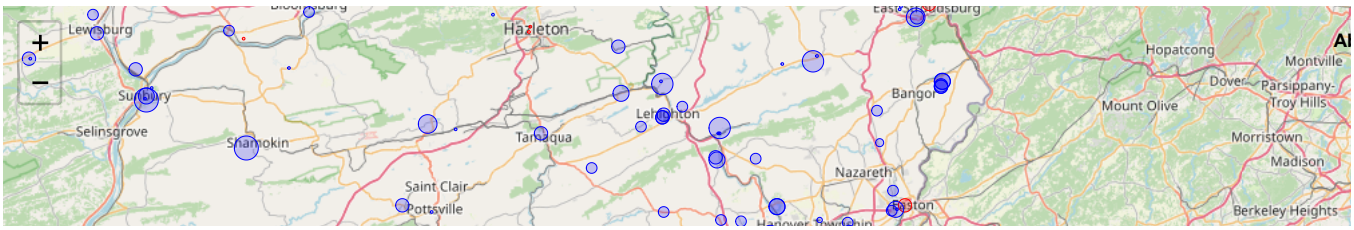
The graph shows the schools which have more than 30% of their enrollees being free lunch recepients. The graphs are based on classifying the data into mostly white (>50% white) and mostly non-white (>50% non-white). The data is filtered using ifelse condition[4]. What is interesting about the two plots is that while the population of non-white schools with high levels of free lunch recepients are few and concentrated in pockets, among white schools, the incidence is very high and spread across the state.

Question 4

Hide

```
#4 - School Attendance

nonwhiteschools = easternschools %>% filter(nonwhite>0.5)
whiteschools = easternschools %>% filter(white>0.5)
attendance = leaflet(districts) %>% addTiles() %>%
    setView(lat=40.2, lng=-75.5,zoom=8.5)%>%
#    addPolygons(weight=1,color="blue",fillOpacity = 0.05) %>%
  addCircleMarkers(data = nonwhiteschools, radius = ~(100-AttendRate), color = "red", weight = 0.75, opacity = 1, fill = TRU
E, fillColor = "red", fillOpacity = 0.2, label = ~city) %>%
  addCircleMarkers(data = whiteschools, radius = ~(100-AttendRate), color = "blue", weight = 0.75, opacity = 1, fill = TRUE,
fillColor = "blue", fillOpacity = 0.2, label = ~city) %>%
  addLegend("topright", colors = c("red", "blue"), labels = c("Mostly Non-White", "Mostly White"), title = "Absence Rate", o
pacity = 1)
attendance
```



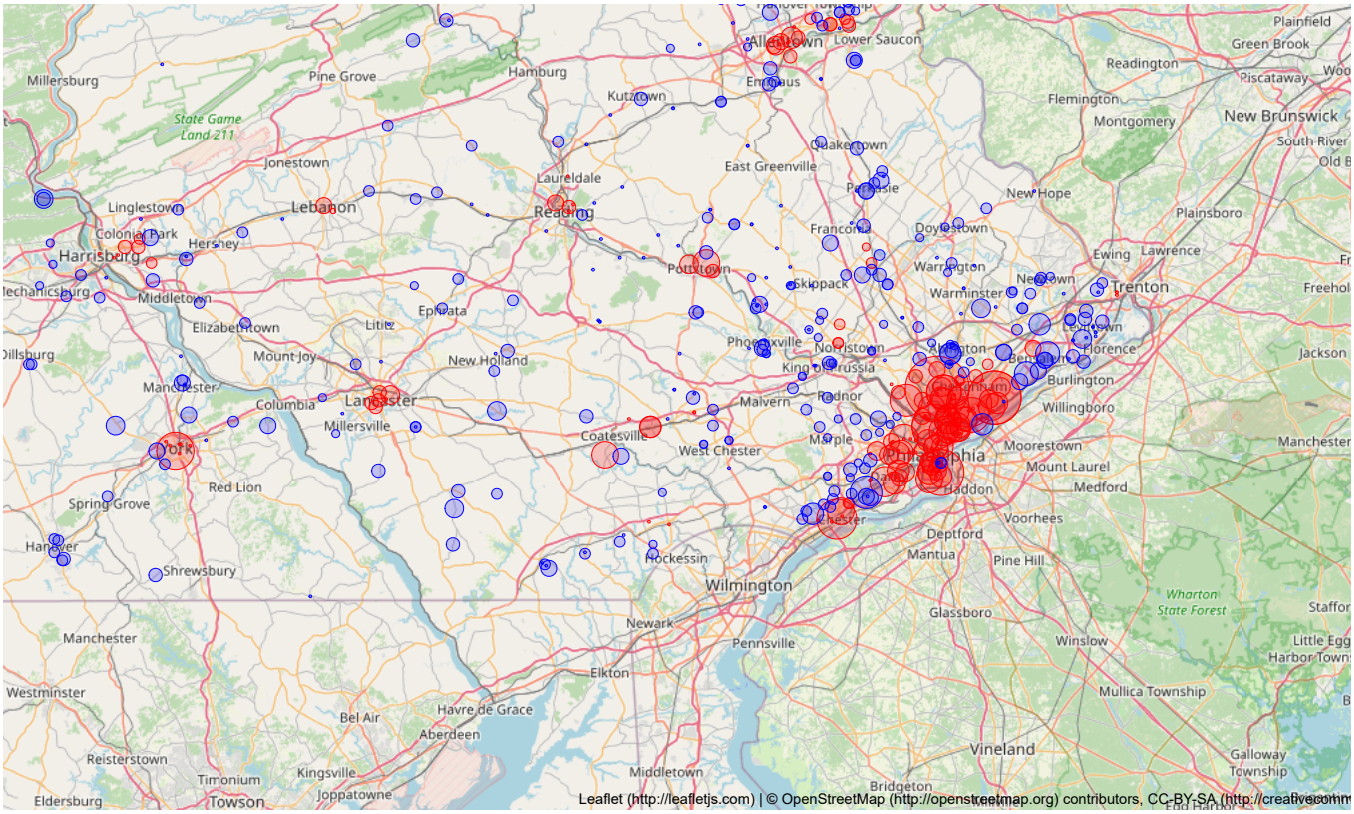


Figure 4: Absence rate among schools by race

The plot shows the absence rate (1-Attendance Rate) of schools split into two groups - schools that are mostly white (>50%) vs mostly non-white. From the plot, we can see that the absence rate among mostly nonwhite schools is much higher than mostly white schools.

References

1. Replaces Occurences Of A Value With Another Value In Set Of Columns Link: https://www.rdocumentation.org/packages/anchors/versions/3.0-8/topics/replace.value (https://www.rdocumentation.org/packages/anchors/versions/3.0-8/topics/replace.value)↩
2. Leaflet for R. GitHub. Link: https://rstudio.github.io/leaflet/ (https://rstudio.github.io/leaflet/)↩
3. Cutting Leaflet Map in R to a specific Area. StackExchange. Link: https://gis.stackexchange.com/questions/231103/cutting-leaflet-map-in-r-to-a-specific-area (https://gis.stackexchange.com/questions/231103/cutting-leaflet-map-in-r-to-a-specific-area)↩
4. R if.else Statement. Data Mentor. Link: https://www.datamentor.io/r-programming/if-else-statement/ (https://www.datamentor.io/r-programming/if-else-statement/)↩