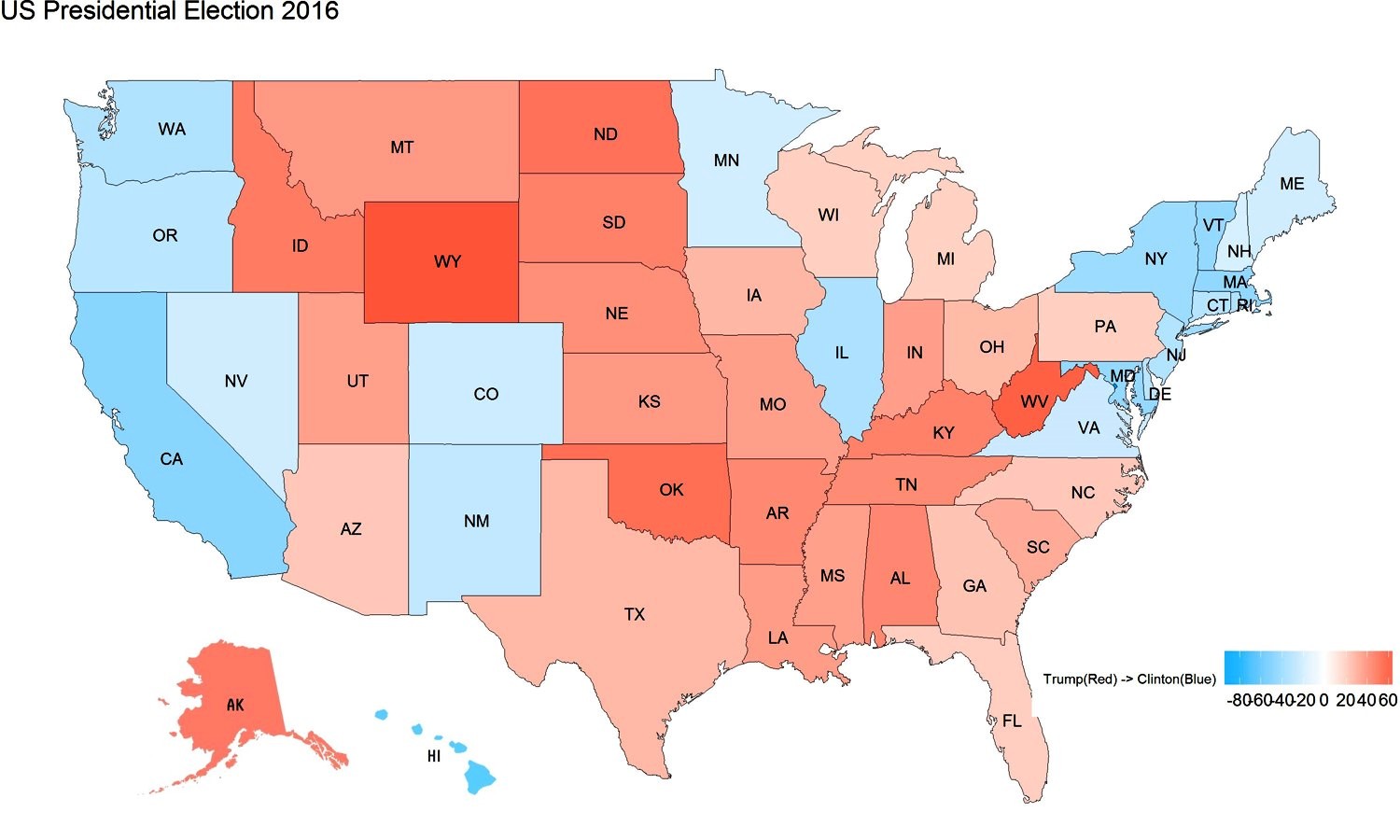
**Mini Project #2**

(a). The project aims to use R to make a map of state-by-state 2016 presidential election results. Following Professor Pankaj Choudhary’s instructions, we have used the following protocol:

**Section 1:**

Step 1: The original election results were saved as us\_2016\_election\_data.csv file, which provide the percentage of votes for Clinton, Trump and other candidates for each state. Following Professor Choudhary’s suggestion, we first extract the percentage difference between Clinton and Trump, which is saved as Candidate = (Trump % - Clinton %) \* 100 in a new file named as us\_2016\_election\_updated\_data.csv. Please note Professor Choudhary specifically instructed that manipulation of the original csv file was allowed for this project.

Step 2: We then wrote a miniproject\_2.R script to plot the election results for the 50 US states (Alaska and Hawaii included). We used the shape file that professor sent it to us to include Alaska and Hawaii on the map. Specifically following professor Choudhary’s instructions, we define “blue” color for the states which voted for Clinton, and “Red” color for the states which voted for Trump. The states which had very close results will be appeared as “white” color. In addition, the darker the blue or red a state is, the bigger the winning of Clinton or Trump is in that state. The final plot was saved as usa\_2016\_election\_results.jpg. Again, it should be emphasized that certain states, such as Michigan and Pennsylvania, appeared unfilled, but actually faithfully reflect tight election results (midpoint).



**Section 2: R code**

Library(sp)

library(raster) # to get map shape file

library(ggplot2) # for plotting and miscellaneous things

library(ggmap) # for plotting

library(plyr) # for merging datasets

library(scales) # to get nice looking legends

library(rgeos) #to add map from a shape file for Alaska and Hawaii

library(rgdal) #to add map from a shape file for Alaska and Hawaii

library(maps)

# Get a shape file of states in the US

usa.df <- map\_data("state")

colnames(usa.df)[5] <- "State"

#print(usa.df)

# Get the data to be plotted

usa.dat <- read.table("C:/Users/nxc161330/My Documents/R/us\_2016\_election\_updated\_data.csv", header = T, sep = ",")

usa.dat$State <- tolower(usa.dat$State)

# Merge the data with the shape file

usa.df <- join(usa.df, usa.dat, by = "State", type = "inner")

# Read data from shape file

ST <- readOGR(dsn=’ cb\_2013\_us\_state\_20m’, "cb\_2013\_us\_state\_20m")

latlong2state <- function(pointsDF) { # Function to include Alaska and Hawaii in the map

states <- ST

# Convert pointsDF to a SpatialPoints object

# USING THE CRS THAT MATCHES THE SHAPEFILE

pointsCRS <- "+proj=longlat +datum=NAD83 +no\_defs +ellps=GRS80 +towgs84=0,0,0"

pointsSP <- SpatialPoints(pointsDF, proj4string=CRS(pointsCRS))

# Use 'over' to get \_indices\_ of the Polygons object containing each point

indices <- over(pointsSP, states)

# Return the state names of the Polygons object containing each point

as.vector(indices$NAME)

}

# Assigning Latitude and Longitude for Alaska and Hawaii

ak <- data.frame(lon = c(-151.0074), lat = c(63.0694))

hi <- data.frame(lon = c(-157.8583), lat = c(21.30694))

# Abbreviations of states and where thy should be plotted

states <- data.frame(state.center, state.abb) # centers of states and abbreviations

subset <- tolower(state.name) #%in% usa.df$State

states <- states[subset, ]

p <- function(data, brks, title) {

ggp <- ggplot() +

geom\_polygon(data = data, aes(x = long, y = lat, group = group, fill = Candidate), color = "black", size = 0.15) +

scale\_fill\_gradient2(low="deepskyblue1",high="firebrick1",midpoint = 0, breaks=brks, space = "Lab", guide="colorbar") +

theme\_nothing(legend = TRUE) +

theme(legend.title=element\_text(size=7),legend.direction ="horizontal", legend.position=c(0.87,0.2)) +

labs(title = title, fill = "Clinton(Blue) -> Trump(Red)") +

geom\_text(data = states, aes(x = x, y = y, label = state.abb), size = 3)

# Drawing Alaska and Hawaii on the map

latlong2state(ak)

latlong2state(hi)

return(ggp)

}

brks.to.use <- seq(-100, 100, by = 20)

figure.title <- "US Presidential Election 2016"

ggsave(p(usa.df, brks.to.use, figure.title), height = 6, width = 5\*2,

file = "C:/Users/nxc161330/My Documents/R/usa\_2016\_election\_results.jpg")

(b). From the plotted map (usa\_2016\_election\_results.jpg), several conclusions could be drawn:

(1). Overall, it is obvious that Trump has carried more states than Clinton, which matches the final election results;

(2). The “blue” states are concentrated in the west coast, as well as the new England state. In contrast, Trump claimed high support in the south, and the Midwest states;

(3). Importantly, the states which supposed to build a “blue” wall, Michigan, Wisconsin and Pennsylvania, are uncharacteristically close to “white” color. This critical fact largely explained the unexpected win of Donald Trump.