

# Class8\_20230330\_DataVisualization\_Apichat

Apichat Photo-A

2023-03-30

```
#load package
library(ggplot2)
library(ggthemes)
library(ggrepel)
library(tidyverse)

## -- Attaching core tidyverse packages -- tidyverse 2.0.0 --
## ✓ dplyr 1.1.0      ✓ readr 2.1.4
## ✓ forcats 1.0.0    ✓ stringr 1.5.0
## ✓ lubridate 1.9.2  ✓ tibble 3.1.8
## ✓ purrr 1.0.1     ✓ tidyr 1.3.0
## --- Conflicts --- tidyverse_conflicts() ---
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag() masks stats::lag()
## ✖ Use the [ ]8;http://conflicted.r-lib.org/[conflicted package]8;;[] to force all conflicts to become errors

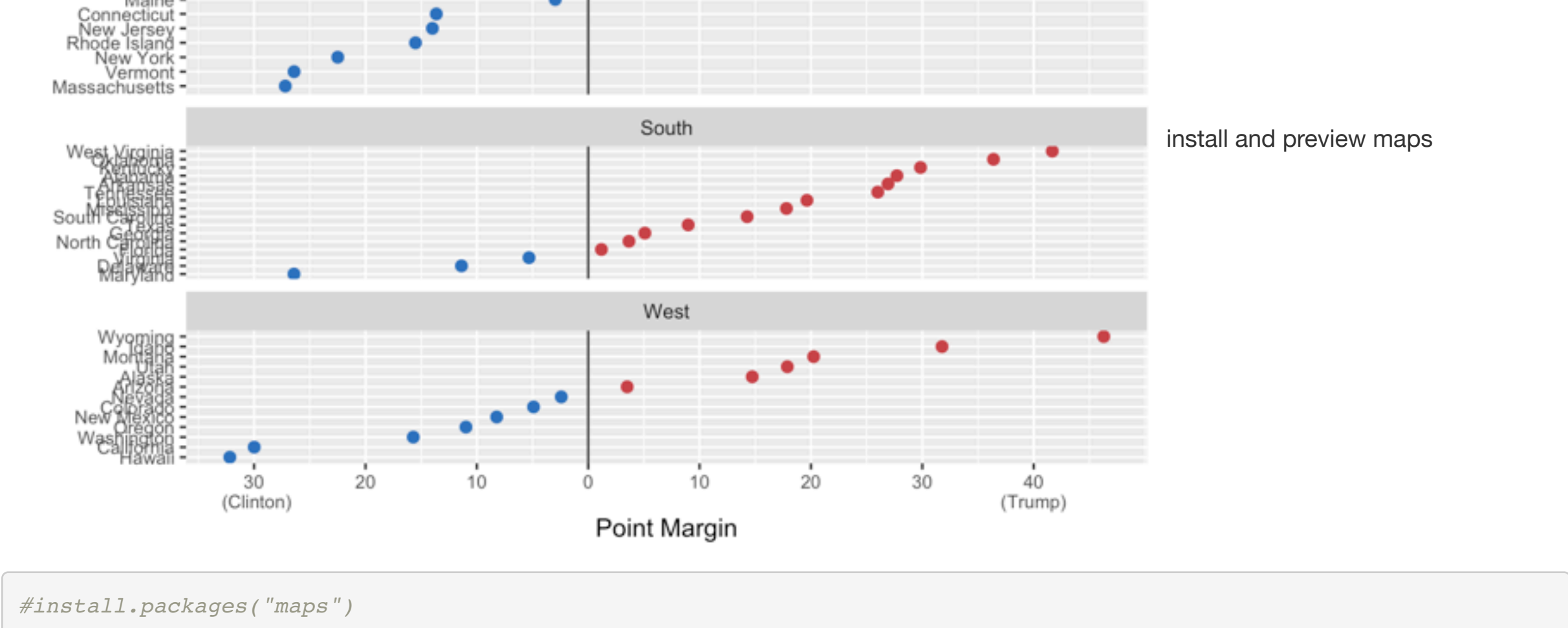
library(socviz)

##review data
election |>
  select(state, total_vote, r_points, pct_trump, party, census) |>
  sample_n(5)

## # A tibble: 5 × 6
##   state total_vote r_points pct_trump party      census
##   <chr>      <dbl>    <dbl>    <dbl> <chr>      <chr>
## 1 Michigan 4824542    0.220    47.2 Republican Midwest
## 2 Oregon    2001336   -11.0     39.1 Democratic West
## 3 Alaska    318608    14.7     51.3 Republican West
## 4 New York  7721795  -22.5     36.5 Democratic Northeast
## 5 Oklahoma  1452992   36.4     65.3 Republican South

party_colors <- c("#2E74C0", "#CB454A")
p0 <- ggplot(data = subset(election, st %nin% "DC"),
             mapping = aes(x = r_points,
                           y = reorder(state, r_points), color = party))
p1 <- p0 + geom_vline(xintercept = 0, color = "gray50") + geom_point(size = 2)
p2 <- p1 + scale_color_manual(values = party_colors)
p3 <- p2 + scale_x_continuous(breaks = c(-30, -20, -10, 0, 10, 20, 30, 40),
                             labels = c("30\n (Clinton)", "20", "10", "0", "10", "20", "30", "40\n (Trump)"))
p3 + facet_wrap(~ census,
                ncol=1, scales="free_y") +
  guides(color=FALSE) + labs(x = "Point Margin", y = "") +
  theme(axis.text=element_text(size=8))

## Warning: The `scale` argument of `guides()` cannot be 'FALSE'. Use 'none' instead as
## of ggplot2 3.3.4.
```



```
#install.packages("maps")
library(maps)

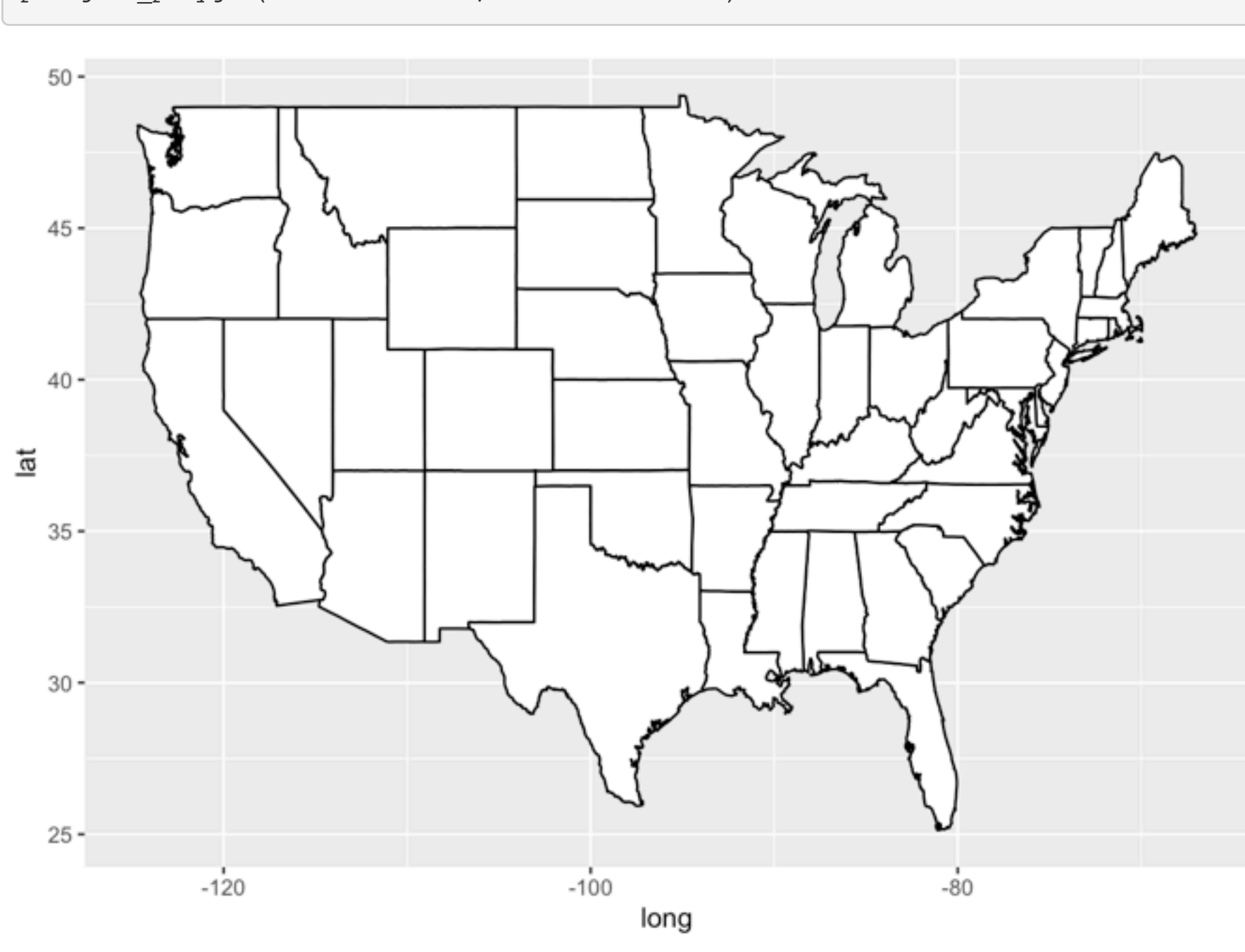
##
## Attaching package: 'maps'

## The following object is masked from 'package:purrr':
##   map

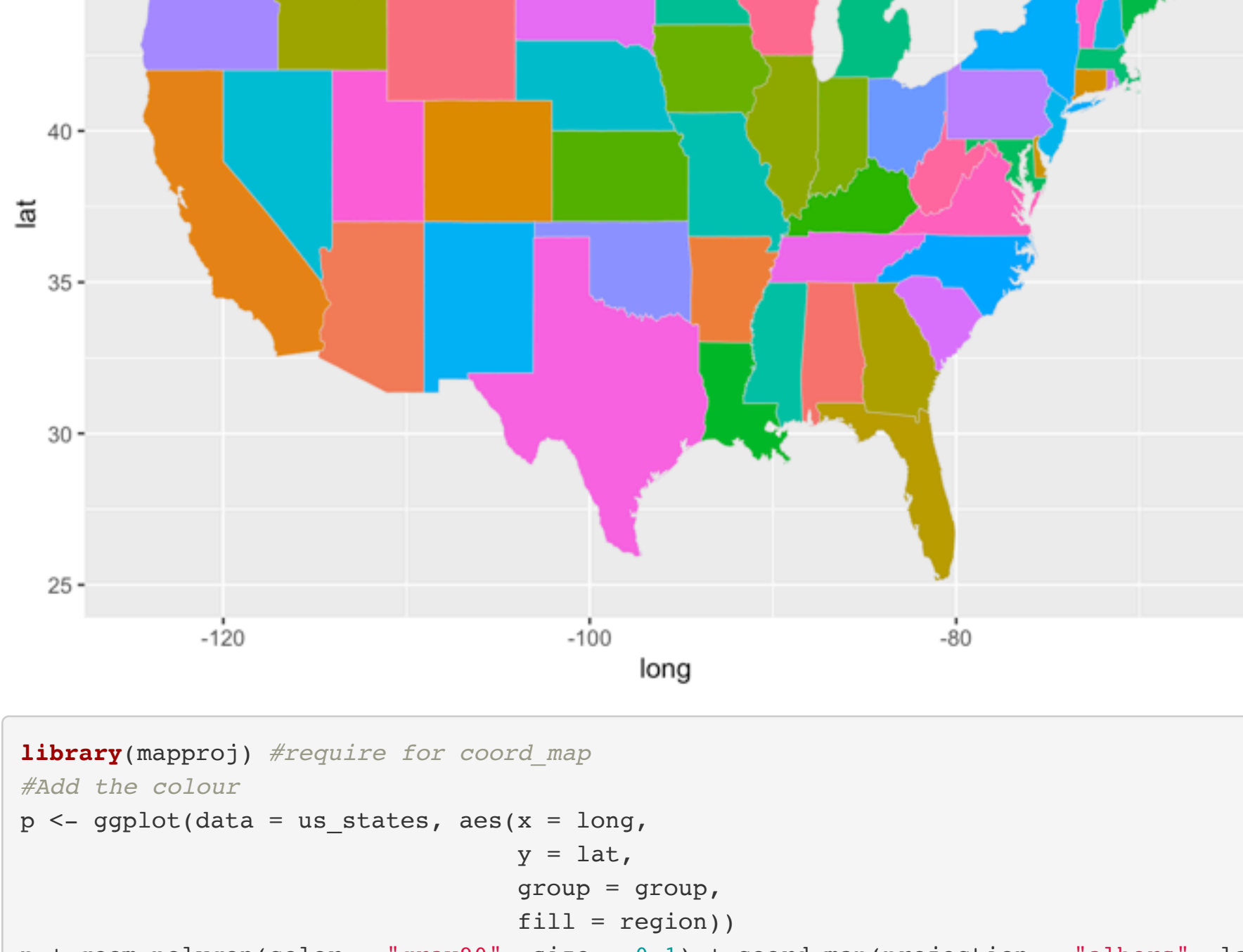
us_states <- map_data("state")
head(us_states)

##      long      lat group order  region subregion
## 1 -87.46201 30.38968     1     1 alabama    <NA>
## 2 -87.48493 30.37249     1     2 alabama    <NA>
## 3 -87.52503 30.37249     1     3 alabama    <NA>
## 4 -87.53076 30.33239     1     4 alabama    <NA>
## 5 -87.57087 30.32665     1     5 alabama    <NA>
## 6 -87.58806 30.32665     1     6 alabama    <NA>

#Create a black and white map
p <- ggplot(data = us_states, mapping = aes(x = long,
                                             y = lat,
                                             group = group))
p + geom_polygon(fill = "white", color = "black")
```

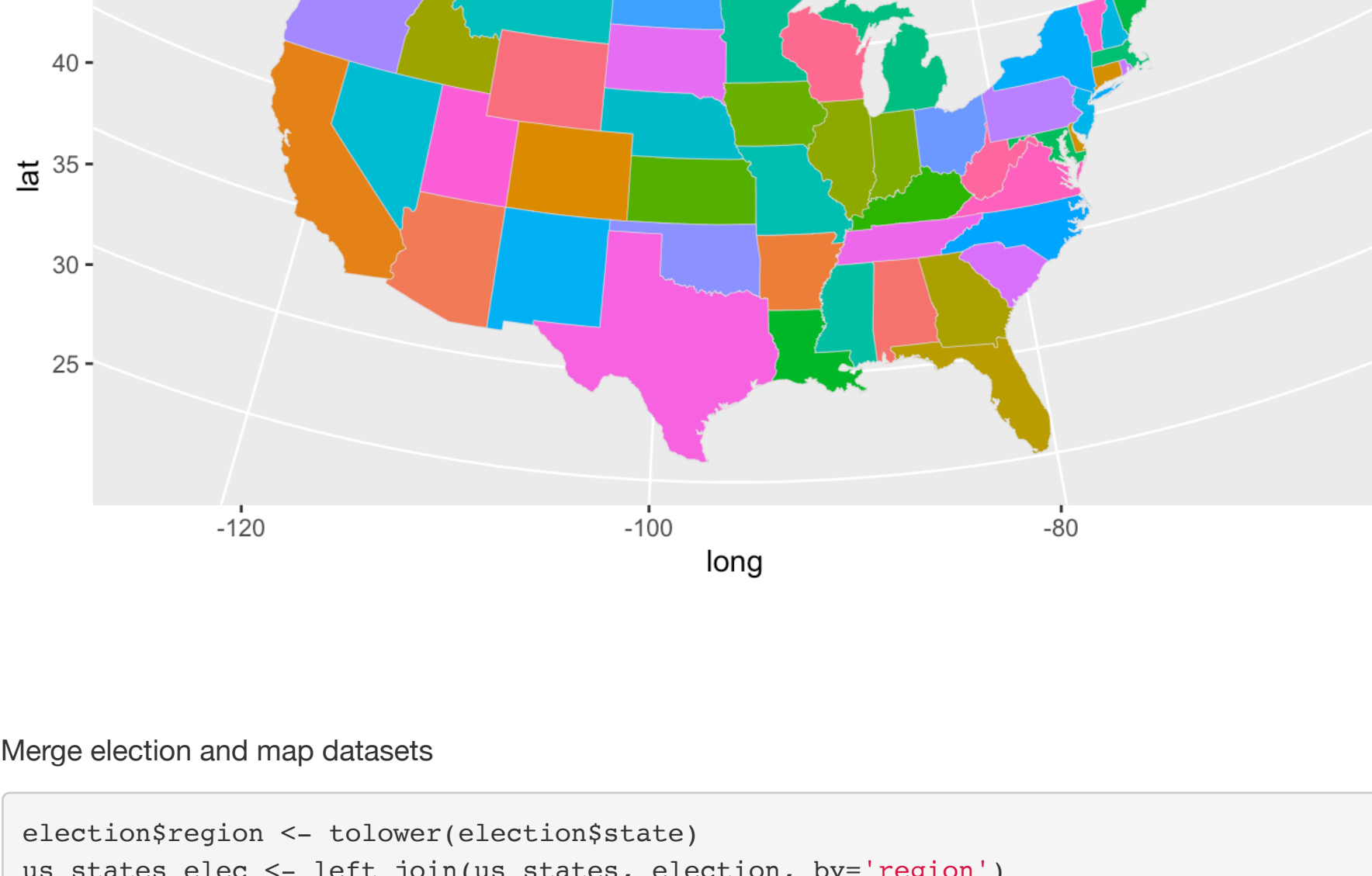


```
#add the colour
p <- ggplot(data = us_states, aes(x = long,
                                  y = lat,
                                  group = group,
                                  fill = region))
p + geom_polygon(color = "gray90", linewidth = 0.1) + guides(fill = FALSE)
```



```
library(mapproj) #require for coord_map
#add the colour
p <- ggplot(data = us_states, aes(x = long,
                                  y = lat,
                                  group = group,
                                  fill = region))
p + geom_polygon(color = "gray90", size = 0.1) + coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  guides(fill = FALSE)
```

## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.  
## i Please use 'linewidth' instead.



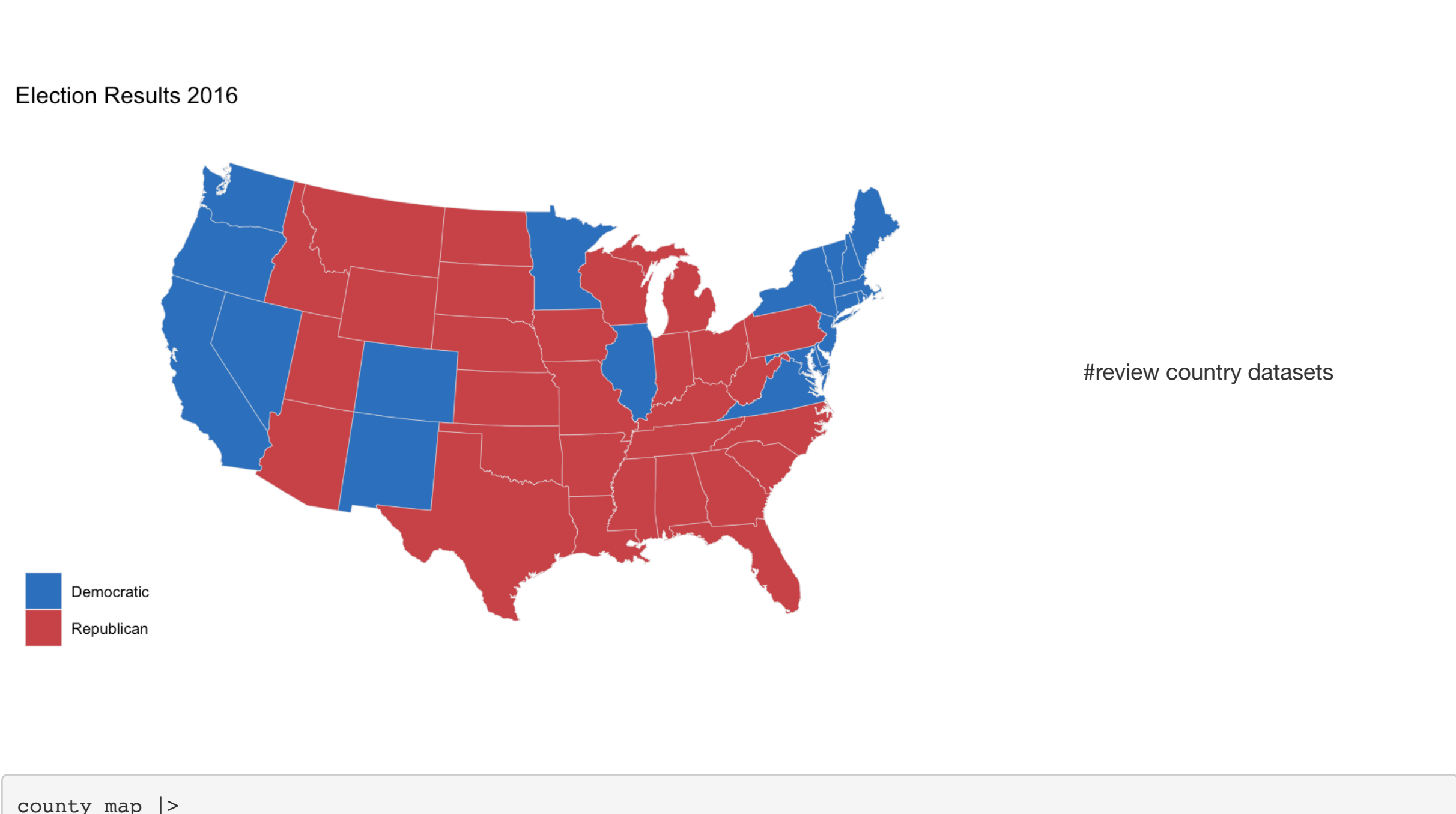
Merge election and map datasets

```
election$region <- tolower(election$state)
us_states_elec <- left_join(us_states, election, by="region")
head(us_states_elec)

##      long      lat group order  region subregion state st fips total_vote
## 1 -87.46201 30.38968     1     1 alabama    <NA> Alabama AL 1 2123372
## 2 -87.48493 30.37249     1     2 alabama    <NA> Alabama AL 1 2123372
## 3 -87.52503 30.37249     1     3 alabama    <NA> Alabama AL 1 2123372
## 4 -87.53076 30.33239     1     4 alabama    <NA> Alabama AL 1 2123372
## 5 -87.57087 30.32665     1     5 alabama    <NA> Alabama AL 1 2123372
## 6 -87.58806 30.32665     1     6 alabama    <NA> Alabama AL 1 2123372
## vote_margin winner party pct_margin r_points d_points pct_clinton
## 1 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## 2 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## 3 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## 4 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## 5 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## 6 588708 Trump Republican 0.2773 27.72 -27.72 34.36
## pct_trump pct_johnson pct_other clinton_vote trump_vote johnson_vote
## 1 62.08 2.09 1.46 729547 1318255 44467
## 2 62.08 2.09 1.46 729547 1318255 44467
## 3 62.08 2.09 1.46 729547 1318255 44467
## 4 62.08 2.09 1.46 729547 1318255 44467
## 5 62.08 2.09 1.46 729547 1318255 44467
## 6 62.08 2.09 1.46 729547 1318255 44467
## other_vote ev_dem ev_rep ev_oth census
## 1 31103 0 9 0 South
## 2 31103 0 9 0 South
## 3 31103 0 9 0 South
## 4 31103 0 9 0 South
## 5 31103 0 9 0 South
## 6 31103 0 9 0 South
```

plot election data on a map

```
party_colors <- c("#2E74C0", "#CB454A")
p0 <- ggplot(data = us_states_elec,
             mapping = aes(x = long,
                           y = lat,
                           group = group,
                           fill = party))
p1 <- p0 + geom_polygon(color = "gray90", size = 0.1) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45)
p2 <- p1 + scale_fill_manual(values = party_colors) +
  labs(title = "Election Results 2016", fill = NULL)
p2 + theme_map()
```



#review country datasets

```
county_map |>
  sample_n(5)

##      long      lat order hole piece group id
## 1 150909.0 299017.78 129260 FALSE 1 05000000US38091.1 38091
## 2 937623.0 -910823.99 103256 FALSE 1 05000000US29155.1 29155
## 3 322397.9 -362124.42 111994 FALSE 1 05000000US31177.1 31177
## 4 2199227.9 40672.03 33431 FALSE 1 05000000US09003.1 09003
## 5 551637.4 -943608.59 19423 FALSE 1 05000000US05015.1 05015
```

```
county_data |>
  select(id, name, state, pop_dens) |>
  sample_n(5)

##      id name state pop_dens
## 1 27095 Mille Lacs County MN ( 100, 50)
## 2 40131 Rogers County OK ( 100, 500)
## 3 21017 Bourbon County KY ( 50, 100)
## 4 22065 Madison Parish LA ( 10, 50)
## 5 31051 Dixon County NE ( 10, 50)
```

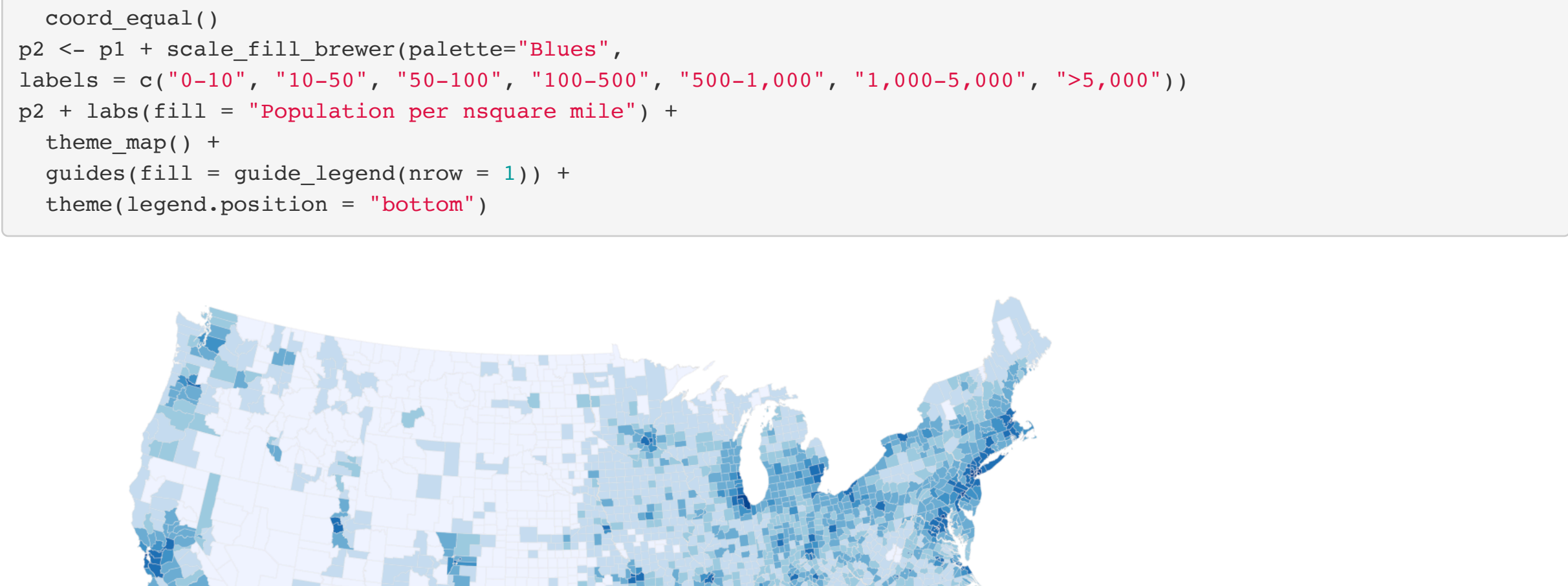
#merge datasets

```
county_full <- left_join(county_map, county_data, by = "id")
head(county_full)
```

```
##      long      lat order hole piece group id name
## 1 1225989 -1275020 1 FALSE 1 05000000US01001.1 01001 Autauga County
## 2 1235424 -1274008 2 FALSE 1 05000000US01001.1 01001 Autauga County
## 3 1243873 -1272331 3 FALSE 1 05000000US01001.1 01001 Autauga County
## 4 1244129 -1267515 4 FALSE 1 05000000US01001.1 01001 Autauga County
## 5 1272010 -1262889 5 FALSE 1 05000000US01001.1 01001 Autauga County
## 6 1276797 -1295514 6 FALSE 1 05000000US01001.1 01001 Autauga County
## state census_region pop_dens pop_dens4 pop_dens6 pct_black pop
## 1 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## 2 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## 3 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## 4 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## 5 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## 6 AL South ( 50, 100) ( 45, 118) ( 82, 215) (15.0,25.0) 55395
## female white black travel_time land_area hh_income su_gun4 su_gun6 fips
## 1 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## 2 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## 3 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## 4 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## 5 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## 6 51.5 78.1 18.4 26.2 594.44 53682 (11,54) (10,12) 1001
## votes_dem_2016 votes_gop_2016 total_votes_2016 per_dem_2016 per_gop_2016
## 1 5908 18110 24661 0.2395685 0.7343579
## 2 5908 18110 24661 0.2395685 0.7343579
## 3 5908 18110 24661 0.2395685 0.7343579
## 4 5908 18110 24661 0.2395685 0.7343579
## 5 5908 18110 24661 0.2395685 0.7343579
## 6 5908 18110 24661 0.2395685 0.7343579
## diff_2016 per_dem_2012 per_gop_2012 diff_2012 winner partywinner16 winner12
## 1 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## 2 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## 3 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## 4 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## 5 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## 6 12202 0.2657577 0.7263374 11012 Trump Republican Romney
## partywinner12 flipped
## 1 Republican No
## 2 Republican No
## 3 Republican No
## 4 Republican No
## 5 Republican No
## 6 Republican No
```

#plotting population density

```
p <- ggplot(data = county_full,
            mapping = aes(x = long,
                          y = lat,
                          fill = pop_dens,
                          group = group))
p1 <- p + geom_polygon(color = "gray90", size = 0.05) +
  coord_equal()
p2 <- p1 + scale_fill_brewer(palette="blues",
                             labels = c("0-10", "10-50", "50-100", "100-500", "500-1,000", "1,000-5,000", ">5,000"))
p2 + labs(fill = "Population per nsquare mile") +
  theme_map() +
  guides(fill = guide_legend(nrow = 1)) +
  theme(legend.position = "bottom")
```



#the range of pop per sq mile

depends on the "pop\_den" range #if we need to change these ranges, me must reclassify before plotting the graph.

#Reference: <https://app.diagrams.net/>

##install

```
#install.packages("Diagrammer")
library(Diagrammer)

#making basic flowchart
grViz(digraph = "digraph flowchart {
  tab1 [label = 'e81', fontname = arial, shape = plaintext, fontcolor = blue]
  tab2 [label = 'e82', fontname = arial, shape = plaintext, fontcolor = blue]
  tab3 [label = 'e83', fontname = arial, shape = plaintext, fontcolor = blue]
  tab4 [label = 'e84', fontname = arial, shape = plaintext]

  tab1 --> tab2 --> tab3 [color = red, arrowhead = vee, arrowsize = 1, penwidth = 5];
  tab2 --> tab4

  [1]: 'Artefact collection in field'
  [2]: 'Preliminary dating of artefacts (visual)'
  [3]: 'Artefacts sent to lab for dating'
  [4]: 'Artefacts put in storage'
  ")
```

