# Congress Data Report

Haoxi Ma

4/18/2020

## Import data and load some packages

```
setwd("/Users/mahaoxi/Desktop/project/Congress data management")
load("Congress99.RData")
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.6.2
```

```
library(mice)
```

```
## Warning: package 'mice' was built under R version 3.6.2
```

```
library(plotrix)
```

```
## Warning: package 'plotrix' was built under R version 3.6.2
```

```
library(maps)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.6.2
```

## Check data property

### 1.Test missing value

```
md.pattern(dat99)
```

```
##  /\     /\
## {  `---'  }
## {  0   0  }
## ==>  V <==  No need for mice. This data set is completely observed.
##  \  \|/  /
##   `-----'
```

```
##        year roll_number issue question result title_or_description vote_result
## 222048    1           1     1        1      1                     1           1
##           0           0     0        0      0                     0           0
##        vote_legislator.party vote_legislator.text vote_legislator.role
## 222048                     1                    1                    1
##                            0                    0                    0
##        vote_legislator.state vote_metadata parsed_date
## 222048                     1             1           1 0
##                            0             0           0 0
```

There is no missing value

**2.Show data structure**

```
str(dat99)
```

```
## 'data.frame':    222048 obs. of  13 variables:
##  $ year                 : int  1999 1999 1999 1999 1999 1999 1999 1999 1999 1999 ...
##  $ roll_number          : int  611 611 611 611 611 611 611 611 611 611 ...
##  $ issue                : Factor w/ 5174 levels "","ADJOURN","H CON RES 100",..: 565 565 565 565 565
##  $ question             : Factor w/ 671 levels "","ON MOTION TO CLOSE PORTIONS OF CONFERENCE",..: 188
##  $ result               : Factor w/ 8 levels "","A","B","F",..: 7 7 7 7 7 7 7 7 7 7 ...
##  $ title_or_description : Factor w/ 11786 levels "","(Targeted Continuing Appropriations) Bulgaria",
##  $ vote_result          : Factor w/ 13 levels "","Aye","Colin Powell",..: 13 13 13 13 13 13 13 13 13
##  $ vote_legislator.party: Factor w/ 4 levels "","D","I","R": 2 2 4 2 2 2 4 4 2 4 2 ...
##  $ vote_legislator.text : Factor w/ 1568 levels "","Abercrombie",..: 2 4 6 13 19 25 27 37 40 43 ...
##  $ vote_legislator.role : Factor w/ 3 levels "","legislator",..: 2 2 2 2 2 2 2 2 2 2 ...
##  $ vote_legislator.state: Factor w/ 52 levels "","AK","AL","AR",..: 12 35 3 22 32 44 44 6 3 48 ...
##  $ vote_metadata        : Factor w/ 16369 levels "","{\"majority\": \"D\", \"congress\": \"101\", \"s
##  $ parsed_date          : Factor w/ 3364 levels "1990-01-23","1990-01-24",..: 1041 1041 1041 1041 104
```

```
names(dat99)
```

```
##  [1] "year"                  "roll_number"           "issue"
##  [4] "question"              "result"                "title_or_description"
##  [7] "vote_result"           "vote_legislator.party" "vote_legislator.text"
## [10] "vote_legislator.role"  "vote_legislator.state" "vote_metadata"
## [13] "parsed_date"
```

There is no missing value but have some "" value and in facotr, this correspond to redundancy levels

## Arrange data in two tables

```
Table1<-dat99[,c("vote_legislator.text","vote_legislator.party","vote_legislator.state",
                 "roll_number")]
names(Table1)<-c("Name","Party","State","roll-number")
Table2<-dat99[,c("roll_number","issue","question","title_or_description","vote_result")]
names(Table2)<-c("roll-number","issue number","question","title/description","votes")
```

## Question1

```
congress.member<-Table1[!duplicated(Table1$Name),]
str(congress.member)#there is "" value in Party and state
```

```
## 'data.frame':    438 obs. of  4 variables:
##  $ Name       : Factor w/ 1568 levels "","Abercrombie",..: 2 4 6 13 19 25 27 37 40 43 ...
##  $ Party      : Factor w/ 4 levels "","D","I","R": 2 2 4 2 2 4 4 2 4 2 ...
##  $ State      : Factor w/ 52 levels "","AK","AL","AR",..: 12 35 3 22 32 44 44 6 3 48 ...
##  $ roll-number: int  611 611 611 611 611 611 611 611 611 611 ...
```

```
levels(congress.member$Party)
```

```
## [1] ""  "D" "I" "R"
```

```
congress.member<-congress.member[congress.member$Party!="",]
congress.member<-congress.member[congress.member$State!="",]
#delete redundancy levels
congress.member$Party<-factor(congress.member$Party)
congress.member$State<-factor(congress.member$State)
```

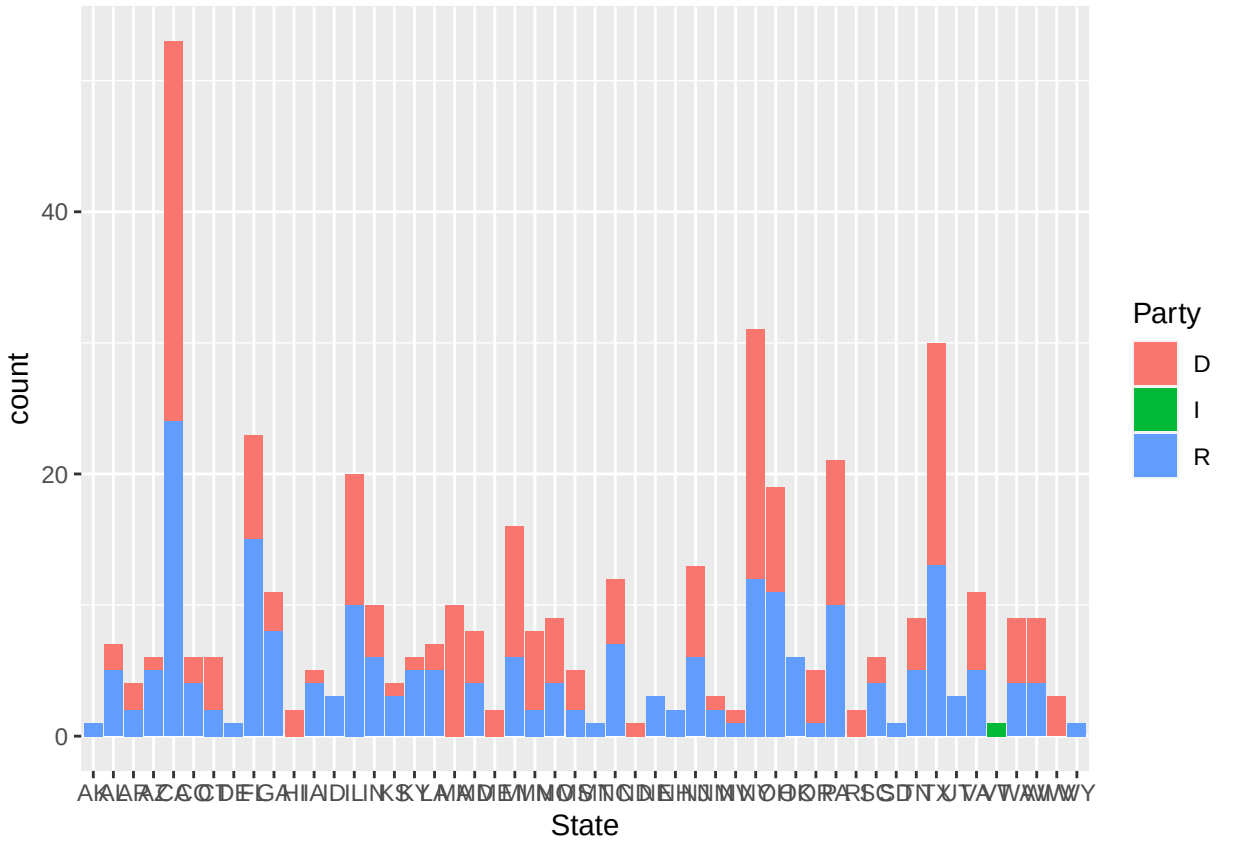
**results**

```
table(congress.member$Party)
```

```
##
##   D   I   R
## 213   1 223
```

## Question2

```
ggplot(data=congress.member,aes(x=State))+geom_bar(aes(fill=Party))
```




```
table(congress.member[c("Party","State")])
```

```
##       State
## Party AK AL AR AZ CA CO CT DE FL GA HI IA ID IL IN KS KY LA MA MD ME MI MN MO
##     D  0  2  2  1 29  2  4  0  8  3  2  1  0 10  4  1  1  2 10  4  2 10  6  5
##     I  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##     R  1  5  2  5 24  4  2  1 15  8  0  4  3 10  6  3  5  5  0  4  0  6  2  4
##       State
## Party MS MT NC ND NE NH NJ NM NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI
##     D  3  0  5  1  0  0  7  1  1 19  8  0  4 11  2  2  0  4 17  0  6  0  5  5
##     I  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0
##     R  2  1  7  0  3  2  6  2  1 12 11  6  1 10  0  4  1  5 13  3  5  0  4  4
##       State
## Party WV WY
##     D  3  0
##     I  0  0
##     R  0  1
```

## Question3

We only consider the Yea and Nay in voting results

```
#function that when same vote_result in R and D, give x=1 else x=0
Sel<-function(x){
```

```
  dat.D<-dat99 %>% filter(roll_number==x) %>% filter(vote_legislator.party=="D")
  dat.R<-dat99 %>% filter(roll_number==x) %>% filter(vote_legislator.party=="R")
  tab.D<-table(dat.D["vote_result"]);tab.R<-table(dat.R["vote_result"])
  ifelse(tab.D["Yea"]>tab.D["Nay"],result.D<-"Agree",result.D<-"Disagree")
  ifelse(tab.R["Yea"]>tab.R["Nay"],result.R<-"Agree",result.R<-"Disagree")
  ifelse(result.D!=result.R,x<-0,x<-1)
}
sketch<-seq(1,611,1);sketch<-unlist(lapply(sketch,Sel));
index<-seq(1,611,1)[sketch==0];
index
```

```
##  [1] 101 103 109 115 116 123 131 132 140 162 163 165 177 194 202 203 210 233 251
## [20] 259 260 264 266 275 298 305 306 307 315 339 348 355 356 366 369 377 378 379
## [39] 387 388 397 403 404 422 423 437 445 447 465 467 468 469 473 480 483 485 502
## [58] 503 504 516 518 527 528 529 547 549 552 558 561 562 564 566 582 590 608 609
```

Now we get the roll number which D and R make different decisions and take #609 as an example

```
dat609.D<-dat99 %>% filter(roll_number==609) %>% filter(vote_legislator.party=="D")
dat609.R<-dat99 %>% filter(roll_number==609) %>% filter(vote_legislator.party=="R")
table(dat609.D["vote_result"])
```

```
##
##                      Aye Colin Powell       Cooper        Lewis          Nay
##            0           0            0            0            0            2
##           No   Not Voting       Pelosi      Present    Ryan (WI) Webster (FL)
##            0            3            0            0            0            0
##          Yea
##          207
```

So there are 207 people in Democratic Party having propensity in #609 roll number

```
table(dat609.R["vote_result"])
```

```
##
##                      Aye Colin Powell       Cooper        Lewis          Nay
##            0           0            0            0            0          217
##           No   Not Voting       Pelosi      Present    Ryan (WI) Webster (FL)
##            0            1            0            0            0            0
##          Yea
##            4
```

So there are 217 people in Republican Party having propensity in #609 roll number

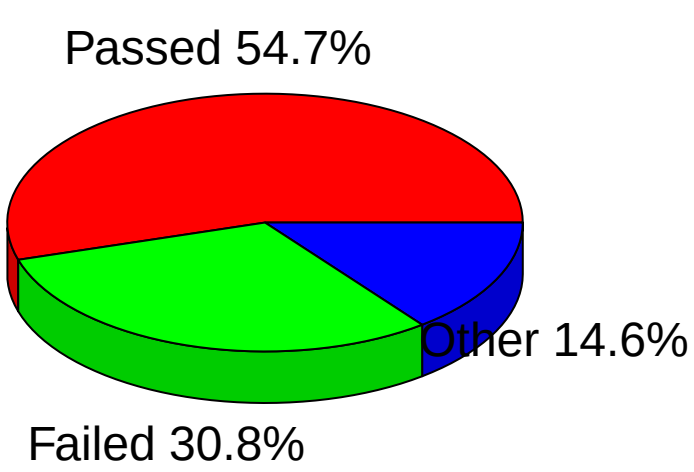
## Question4

### (1)

```
levels(dat99$result)
```

```
## [1] ""  "A" "B" "F" "G" "H" "P" "R"
```

```
res<-dat99[dat99$result!="",]
res$result<-factor(res$result)
res<-res[!duplicated(res$roll_number),]
tabular<-as.data.frame(table(res$result))
num_P<-tabular$Freq[which(tabular$Var1=="P")]
num_F<-tabular$Freq[which(tabular$Var1=="F")]
num_other<-sum(tabular$Freq[which(tabular$Var1!="P"&tabular$Var1!="F")])
number<-c(num_P,num_F,num_other);Name<-c("Passed","Failed","Other")
pct<-round(number/sum(number)*100,digit=1)
lbls<-paste(Name,pct) %>% paste("%",sep="")
pie3D(number,labels=lbls,col=rainbow(length(lbls)),main="Vote Results",theta = pi/4)
```




```
table(sketch)
```

```
## sketch
##   0   1
##  76 535
```

So #76 votes were determined by party affiliations

## Question5

```
Table2<-Table2[which(Table2$`issue number`!="" & Table2$votes!=""),]
Table2$`issue number`<-factor(Table2$`issue number`)
Table2$votes<-factor(Table2$votes)
iss.votes<-as.data.frame.matrix(table(Table2[c("issue number","votes")]))
iss.votes
```

```
##                 Aye  Nay   No Not Voting Present  Yea
## ADJOURN         219 1853 1022        211       3  161
## H CON RES 102     0    0    0         10       0  423
## H CON RES 107     0    0    0         66      13  355
## H CON RES 117     0    5    0         64       0  365
## H CON RES 121     0    0    0         50       2  381
## H CON RES 140     0    1    0         31       1  400
## H CON RES 158     0    0    0         17       0  417
## H CON RES 165     0   12    0         17       0  404
## H CON RES 180     0   41    0         10      72  311
## H CON RES 184   396    0    0         37       0    0
## H CON RES 187     0    2    0         29       0  402
## H CON RES 188     0    0    0          9       0  424
## H CON RES 190     0    1    0          9       0  423
## H CON RES 194     0    0    0         42       0  391
## H CON RES 196     0    0    0         31       0  402
## H CON RES 206     0    4    0         22       0  407
## H CON RES 208     0   48    0         11       3  371
## H CON RES 211     0    4    0         33       0  396
## H CON RES 213     0    3    0         19       0  411
## H CON RES 214     0  193    0         14       2  224
## H CON RES 222     0    0    0         34       0  399
## H CON RES 223     0    0    0         16       0  417
## H CON RES 75      0    1    0         16       1  416
## H CON RES 82      0  290    0          4       0  139
## H CON RES 84      0    2    0         17       1  413
## H CON RES 88      0   13    0         19       4  397
## H CON RES 94      0  140    0          8      11  275
## H J RES 33      305  310  124         14       0  115
## H J RES 44        0  427    0          5       0    2
## H J RES 57        0  260    0          3       1  170
## H J RES 58        0  297    0          6       0  130
## H J RES 65        0    0    0         11       0  422
## H J RES 68        0    2    0          9       1  421
## H J RES 71        0    2    0         11       0  421
## H J RES 73        0    2    0          7       0  424
## H J RES 75        0    6    0         10       0  417
## H J RES 80        0    8    0         23       0  403
## H J RES 82      403  420   16         28       0    1
## H R 10         2063  232 1343         72       0  198
## H R 100           0    0    0         65       0  368
## H R 1000        678    0  603         22       0    0
## H R 1033          0    1    0         51       0  381
## H R 1074        446    0  374         46       0    0
## H R 1141          0  447    0         20       1  832
```

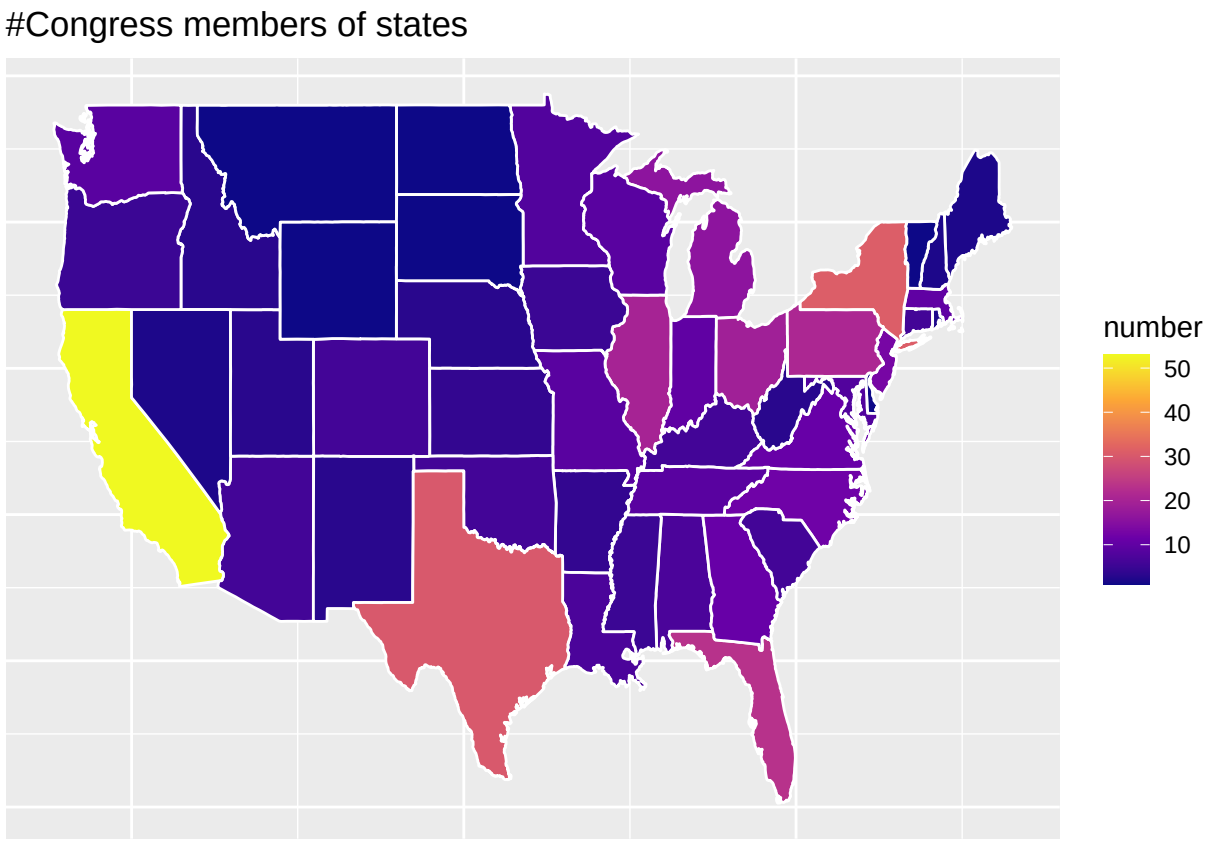
```
## H R 1175         0    5    0         25       1  836
## H R 1180         0   11    0         27       0  830
## H R 1218       270  268  159          7       0  164
## H R 1219         0    0    0         17       0  416
## H R 1251         0    0    0         71       0  362
## H R 1259         0  234    0         12       0  621
## H R 1400         0    1    0        101       0  332
## H R 1401      3119    0 1979        112       0    0
## H R 1402      1000    0 1981         50       0    0
## H R 1431         0  106    0         17       1  309
## H R 1451         0    2    0         20       0  411
## H R 1477         0    1    0         49       0  383
## H R 1480         0    5    0         11       0  418
## H R 1487         0    2    0         23       0  408
## H R 150          0    0    0         14       0  420
## H R 1501      6326 1115 2638        308       0 2194
## H R 1550         0    3    0         13       0  417
## H R 1554         0    8    0         14       0  411
## H R 1555        68    0  343         22       0    0
## H R 1569       249    0  180          5       0    0
## H R 1654       947    0 1622         30       0    0
## H R 1658       530    0  316         22       0    0
## H R 1663         0    0    0          9       0  424
## H R 1664       628  105 1059         64       1  311
## H R 1691       306  234  118         20       0  190
## H R 1714       900  122  346        130       0  234
## H R 1802         0    6    0         48       0  380
## H R 1833         0    2    0         21       0  410
## H R 1875       923    0 1216         26       0    0
## H R 1883         0    0    0         14       0  419
## H R 1887         0   42    0         19       0  372
## H R 1905       130  460  263        105       1  779
## H R 1906      2178  586 4160        204       3  673
## H R 1907         0   43    0         14       0  376
## H R 1915         0    4    0         59       0  370
## H R 1993      1352    0  785         28       0    0
## H R 1995      1476    0  645         45       0    0
## H R 2         1799    0 1164         68       0    0
## H R 202          0    5    0         23       0  405
## H R 2031         0  211    0         20       0  635
## H R 2084       530   94  331         47       1  733
## H R 2116         0   46    0         18       0  369
## H R 2122      2809    0 1871         91       5    0
## H R 2130         0    1    0          9       0  423
## H R 2140         0    9    0         30       0  394
## H R 2260       619    0  663         17       0    0
## H R 2280         0    0    0         10       0  424
## H R 2300       597    0  660         44       0    0
## H R 2303         0    7    0         38       0  388
## H R 2336       183    0  231         19       0    0
## H R 2389       460    0  394         12       0    0
## H R 2415      3238    0 1833        119       6    0
## H R 2436       470  172  382         21       0  254
## H R 2465         0   12    0         26       0  830
```

```
## H R 2466      3044  446 3299        205       0  820
## H R 2488       822  640  889         54       0  631
## H R 2490       679  694  992        109       0  563
## H R 2506         0    7    0          9       0  417
## H R 2561         0  107    0         55       0 1139
## H R 2587       668  298  616         45       0  541
## H R 2605       609   95  246         36       0  747
## H R 2606      1604  246 2203        114       0  599
## H R 2670      2484  423 2610        116       0  432
## H R 2679         0    5    0         13       0  415
## H R 2684      1811  318 2397        159       0  947
## H R 2723       773    0  942         21       0    0
## H R 2737         0    0    0         78       0  355
## H R 2886         0    0    0         29       0  404
## H R 2904         0    1    0         46       0  386
## H R 2910         0    4    0          9       0  420
## H R 2990       211  372  220         13       0  484
## H R 3064         0  833    0         28       1  440
## H R 3073       532   93  737         42       0  328
## H R 3075         0   25    0         20       0  388
## H R 3085         0  419    0          9       5    0
## H R 3164         0   26    0         22       0  385
## H R 3194         0  564    0         16       0  724
## H R 3196         0  158    0         41       0  667
## H R 3257         0    0    0         32       0  401
## H R 348          0    4    0         80       0  349
## H R 4            0   71    0         18       0  345
## H R 417       1837  177 2852         80       0  252
## H R 434          0  163    0         37       0  234
## H R 435          0    1    0         57       0  375
## H R 468          0    2    0         21       0  410
## H R 658        396    0    6         31       0    0
## H R 659          0    4    0         12       0  418
## H R 747          0    0    0         17       0  416
## H R 754          0    2    0         41       0  390
## H R 764       1238    0   34         29       0    0
## H R 775        982   24 1151         50       0  830
## H R 833       1041  108 1065         72       0  313
## H R 883        860    0  404         35       0    0
## H R 987          0  209    0          8       0  217
## H RES 157        0    1    0         24       0  408
## H RES 158        0  190    0         16       0  227
## H RES 159        0  171    0         10       0  253
## H RES 165        0    0    0         13       0  420
## H RES 166        0  188    0          9       0  236
## H RES 169        0    1    0         20       0  412
## H RES 173        0  109    0          9       0  315
## H RES 178        0    0    0         15       0  418
## H RES 180        0  178    0         15       0  240
## H RES 181        0    0    0         20       0  413
## H RES 185        0   10    0         21       0  402
## H RES 186        0  205    0          6       0  223
## H RES 190      448  198  376         68       0  213
## H RES 200        0   75    0          6       0  354
```

```
## H RES 209        0  189    0          6       0  240
## H RES 218        0    3    0         15       0  416
## H RES 226        0    0    0          8       1  425
## H RES 234        0    1    0         10       0  423
## H RES 235        0  203    0          5       0  227
## H RES 246        0  147    0         12       0  276
## H RES 250        0  141    0         31       0  263
## H RES 253        0  187    0         19       0  227
## H RES 256      219    0  208          7       0    0
## H RES 260        0  201    0          5       0  227
## H RES 263        0  172    0          6       0  256
## H RES 273        0  205    0          8       0  221
## H RES 274        0  203    0          7       0  224
## H RES 275        0  208    0          8       0  217
## H RES 281        0  172    0          8       0  253
## H RES 292        0   38    0          4       1  390
## H RES 295        0  181    0         11       0  241
## H RES 297        0    0    0          9       0  424
## H RES 303        0    5    0          7       0  421
## H RES 306        0    2    0          8       6  417
## H RES 308        0    1    0         10       0  422
## H RES 312        0    0    0         13       0  420
## H RES 317        0  188    0         16       0  230
## H RES 322        0    0    0         15       1  417
## H RES 323        0  209    0          4       0  221
## H RES 325        0    0    0         19       0  414
## H RES 330        0  202    0         14       0  217
## H RES 335      221  204  204         16       0  221
## H RES 337        0  196    0          9       0  228
## H RES 338        0  201    0         19       0  214
## H RES 344        0    0    0         44       0  389
## H RES 345        0  206    0          7       0  221
## H RES 349        0    0    0         24       0  409
## H RES 353        0  200    0         10       1  222
## H RES 355      335    0   79         20       0    0
## H RES 367        0  144    0         11       0  278
## H RES 374        0  202    0         17       0  214
## H RES 385      668   45  164         47       0  375
## H RES 386        0  204    0          4       0  226
## H RES 59         0  133    0         21       1  278
## H RES 62         0    1    0         18       1  414
## H RES 94         0    0    0         51       0  382
## JOURNAL          0 1294    0        706      22 8376
## MOTION        1609  204 1297        141       0  218
## QUORUM           0    0    0         34     399    0
## S 1059         375  452   45         45       5  813
## S 249            0    1    0         18       0  414
## S 293            0    1    0         26       0  406
## S 440            0  291    0         14       0  128
## S 800            0    2    0          7       0  424
## S 900            0  189    0         76       0  603
## S CON RES 21     0  213    0          8       0  213
## S CON RES 35     0  178    0          7       0  249
```

### Draw a plot for states corresponding to #congress member

```
member<-as.data.frame(table(congress.member$State))
names(member)<-c("region","number")
#order the region
member$region<-factor(member$region,levels=state.abb)
levels(member$region)<-state.name
member$region<-tolower(as.character(member$region))
map_state<-map_data("state")
num.member<-left_join(map_state,member,by="region")
mhx<-ggplot(num.member,aes(long,lat,group=group))+geom_polygon(aes(fill=number),color="white")+
      scale_fill_viridis_c(option="C")+labs(title="#Congress members of states")+
      theme(axis.text = element_blank(),axis.ticks = element_blank(),
            axis.title = element_blank())
mhx
```




```
rm(list=ls())
```