Perception of climate change in the US

2022-10-22

Using packages and inspecting data

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(readxl)
library(ellipse)
```

```
##
## Attaching package: 'ellipse'
```

```
## The following object is masked from 'package:graphics':
##
##     pairs
```

```
df <- read.csv("data.csv")
head(df)
```

```
##    GeoType GEOID    GeoName  TotalPop happening happeningOppose  human
## 1 National  9999         US 251268391    71.827          14.459 56.540
## 2    State     1    Alabama   3779877    63.291          17.818 49.350
## 3    State     2     Alaska    552673    70.188          17.444 53.277
## 4    State     4    Arizona   5414955    72.246          15.191 56.428
## 5    State     5   Arkansas   2295104    66.366          18.109 51.238
## 6    State     6 California  30261354    77.466          10.708 63.349
##   humanOppose consensus consensusOppose affectweather affectweatherOppose
## 1      30.367    57.303          23.389        64.469               7.183
## 2      36.640    47.468          26.865        53.640               8.653
## 3      33.621    56.628          26.673        62.254               7.558
## 4      30.939    57.135          24.560        64.224               7.578
## 5      35.455    48.541          27.555        55.891               8.619
## 6      25.064    63.498          19.547        70.137               5.759
##   worried worriedOppose harmplants harmplantsOppose futuregen futuregenOppose
## 1  65.228        34.314     71.335           19.210    70.640          19.184
## 2  57.555        42.345     62.967           23.185    61.855          23.795
## 3  63.234        36.563     71.375           21.908    69.202          22.228
## 4  65.733        34.264     70.234           20.205    69.209          20.444
```

```
## 5  59.382        40.523     66.879           22.437    65.584          22.672
## 6  72.696        27.195     77.310           14.498    76.511          14.800
##   devharm devharmOppose harmUS harmUSOppose personal personalOppose timing
## 1  67.893        20.765 64.452       26.617   46.770         44.560 59.197
## 2  61.125        24.395 57.374       30.442   40.623         47.664 52.176
## 3  67.506        23.561 63.099       30.039   43.200         49.577 56.102
## 4  67.539        21.605 65.219       27.115   47.802         44.501 58.343
## 5  63.907        24.238 59.569       30.270   40.614         48.492 53.008
## 6  74.256        16.033 70.941       21.472   54.317         38.341 65.554
##   timingOppose    exp expOppose fundrenewables fundrenewablesOppose regulate
## 1       40.359 45.778    53.512         76.943                20.985   71.868
## 2       47.824 38.020    61.126         72.195                26.033   67.147
## 3       43.898 41.390    57.274         76.889                21.068   69.143
## 4       41.657 46.519    53.392         76.866                21.141   71.210
## 5       46.992 38.579    60.084         72.457                25.698   67.113
## 6       34.446 53.363    46.232         79.019                18.983   74.122
##   regulateOppose CO2limits CO2limitsOppose reducetax reducetaxOppose supportRPS
## 1         25.563    65.779          33.302    65.831          32.352     64.326
## 2         30.456    59.434          39.972    60.923          36.501     57.439
## 3         28.313    61.738          37.568    63.349          34.374     58.291
## 4         26.770    64.699          34.800    62.906          34.452     63.837
## 5         29.990    61.387          38.089    62.730          34.906     59.711
## 6         23.518    70.871          28.714    71.065          27.816     68.414
##   supportRPSOppose rebates rebatesOppose drillANWR drillANWROppose
## 1           34.187  77.404        20.365    29.839          68.704
## 2           42.056  74.698        23.258    34.690          63.626
## 3           40.420  77.640        20.194    29.699          68.770
## 4           35.114  76.897        21.195    33.418          65.079
## 5           39.241  73.967        23.233    33.624          64.056
## 6           30.325  79.744        18.224    26.211          72.202
##   drilloffshore drilloffshoreOppose teachGW teachGWOppose corporations
## 1        48.890              49.478  77.456        21.979       69.912
## 2        61.394              36.878  73.858        26.035       64.323
## 3        48.399              50.435  76.014        23.882       68.999
## 4        52.800              45.077  77.132        22.854       69.091
## 5        55.773              41.699  74.951        24.553       66.064
## 6        39.442              59.169  81.154        18.804       73.172
##   corporationsOppose president presidentOppose congress congressOppose governor
## 1             11.077    52.495          26.677   61.187         21.938   56.560
## 2             12.140    45.927          31.626   54.046         26.504   52.885
## 3             10.717    48.096          32.170   57.283         26.207   49.134
## 4             11.214    51.500          28.109   60.166         23.652   55.131
## 5             11.812    47.695          33.444   56.360         27.164   53.158
## 6             10.685    58.158          23.047   65.822         19.832   56.784
##   governorOppose localofficials localofficialsOppose citizens citizensOppose
## 1         17.557         58.752               16.373   64.899         11.981
## 2         18.306         54.742               16.873   59.194         13.723
## 3         20.008         55.131               17.868   63.117         11.010
## 4         17.971         57.356               17.500   63.935         12.103
## 5         18.019         54.652               17.396   60.609         13.029
## 6         20.279         61.893               16.243   69.011         10.978
##   discuss discussOppose mediaweekly mediaweeklyOppose priority priorityOppose
## 1  35.447        64.041      32.722            65.987   55.257         23.707
## 2  27.300        72.726      25.779            73.339   47.646         29.602
```

```
## 3  39.192        59.980      35.095            63.939   50.147         27.587
## 4  37.305        62.536      35.705            63.332   55.122         26.676
## 5  29.337        70.436      26.537            72.519   47.375         29.015
## 6  42.648        57.208      37.013            61.839   63.704         19.064
```

```
dim(df)
```

```
## [1] 4563   64
```

```
unique(df$GeoType)
```

```
## [1] "National" "State"    "County"   "cd116"    "CBSA"
```

We are only interested in National, State and CBSA level (i.e. metro area with at least 10,000 people), thus can filter the other rows.

```
interested_geo <- c("State", "National", "CBSA")
clean_df <- df[df$GeoType %in% interested_geo,]
clean_df$GeoType <- as.factor(clean_df$GeoType)

dim(clean_df)
```

```
## [1] 985  64
```

```
unique(clean_df$GeoType)
```

```
## [1] National State    CBSA
## Levels: CBSA National State
```

Corregram plot

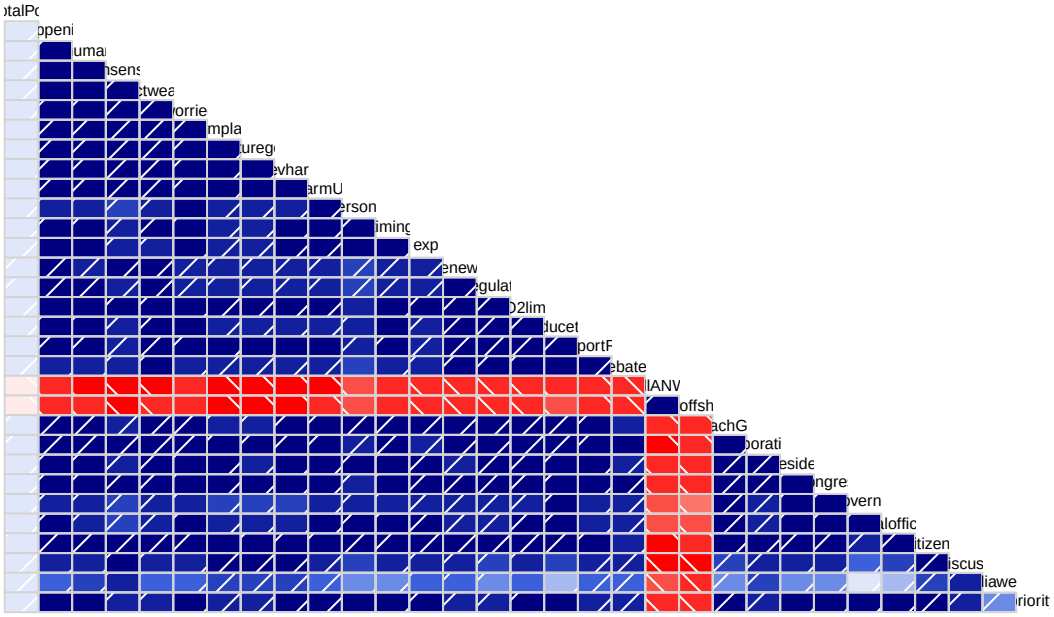
```
##    TotalPop happening  human consensus affectweather worried harmplants
## 1 251268391    71.827 56.540    57.303        64.469  65.228     71.335
## 2   3779877    63.291 49.350    47.468        53.640  57.555     62.967
## 3    552673    70.188 53.277    56.628        62.254  63.234     71.375
## 4   5414955    72.246 56.428    57.135        64.224  65.733     70.234
## 5   2295104    66.366 51.238    48.541        55.891  59.382     66.879
## 6  30261354    77.466 63.349    63.498        70.137  72.696     77.310
##   futuregen devharm harmUS personal timing    exp fundrenewables regulate
## 1    70.640  67.893 64.452   46.770 59.197 45.778         76.943   71.868
## 2    61.855  61.125 57.374   40.623 52.176 38.020         72.195   67.147
## 3    69.202  67.506 63.099   43.200 56.102 41.390         76.889   69.143
## 4    69.209  67.539 65.219   47.802 58.343 46.519         76.866   71.210
## 5    65.584  63.907 59.569   40.614 53.008 38.579         72.457   67.113
## 6    76.511  74.256 70.941   54.317 65.554 53.363         79.019   74.122
##   CO2limits reducetax supportRPS rebates drillANWR drilloffshore teachGW
## 1    65.779    65.831     64.326  77.404    29.839        48.890  77.456
## 2    59.434    60.923     57.439  74.698    34.690        61.394  73.858
## 3    61.738    63.349     58.291  77.640    29.699        48.399  76.014
## 4    64.699    62.906     63.837  76.897    33.418        52.800  77.132
## 5    61.387    62.730     59.711  73.967    33.624        55.773  74.951
## 6    70.871    71.065     68.414  79.744    26.211        39.442  81.154
##   corporations president congress governor localofficials citizens discuss
## 1       69.912    52.495   61.187   56.560         58.752   64.899  35.447
## 2       64.323    45.927   54.046   52.885         54.742   59.194  27.300
## 3       68.999    48.096   57.283   49.134         55.131   63.117  39.192
## 4       69.091    51.500   60.166   55.131         57.356   63.935  37.305
## 5       66.064    47.695   56.360   53.158         54.652   60.609  29.337
```

```
## 6       73.172    58.158   65.822   56.784         61.893   69.011  42.648
##   mediaweekly priority
## 1      32.722   55.257
## 2      25.779   47.646
## 3      35.095   50.147
## 4      35.705   55.122
## 5      26.537   47.375
## 6      37.013   63.704
```

```
library(corrgram)

corrgram(R, order = NULL, lower.panel = panel.shade, upper.panel = NULL, text.panel = panel.txt,
    main = "Predictors of Climate Change Behavior Corrgram Plot")
```

Predictors of Climate Change Behavior Corrgram Plot

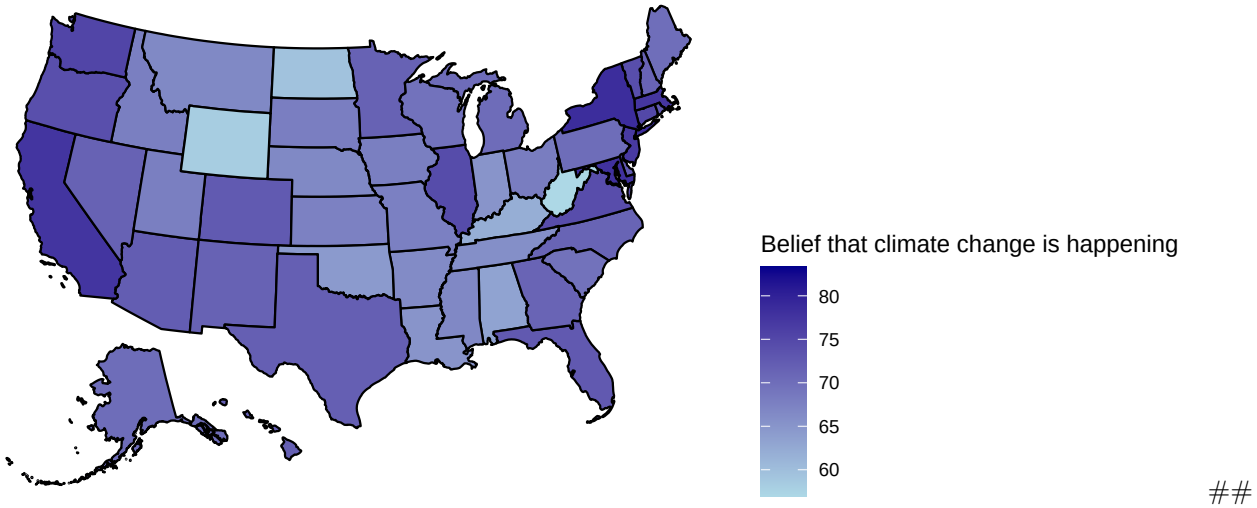


State areas

```
state <- clean_df[clean_df$GeoType == 'State',]
names(state)[names(state) == 'GeoName'] <- 'state'
chosen_vars <- c('GEOID', 'state', 'TotalPop', 'happening', 'human', 'worried')
state <- state[, chosen_vars]
head(state)
```

```
##   GEOID      state TotalPop happening  human worried
## 2     1    Alabama  3779877    63.291 49.350  57.555
## 3     2     Alaska   552673    70.188 53.277  63.234
## 4     4    Arizona  5414955    72.246 56.428  65.733
## 5     5   Arkansas  2295104    66.366 51.238  59.382
## 6     6 California 30261354    77.466 63.349  72.696
## 7     8   Colorado  4349342    72.559 57.099  65.978
```

```
library(usmap)
library(ggplot2)
plot_usmap(data=state, values="happening", color="black") +
  scale_fill_continuous(low="lightblue", high="darkblue", name='Belief that climate change is happening
  theme(legend.position = "right")
```



##

Metro areas

```
metro <- clean_df[clean_df$GeoType == 'CBSA',] %>% select(-contains(c('Oppose')))
head(metro)
```

```
##      GeoType GEOID      GeoName TotalPop happening  human consensus
## 3631    CBSA 10100 Aberdeen, SD    32710    67.750 51.553    54.039
## 3632    CBSA 10140 Aberdeen, WA    57845    68.772 53.170    52.291
## 3633    CBSA 10180  Abilene, TX   130201    62.129 48.285    46.597
## 3634    CBSA 10220      Ada, OK    29357    65.013 49.048    44.626
## 3635    CBSA 10300   Adrian, MI    77534    63.537 51.129    50.364
## 3636    CBSA 10420    Akron, OH   558902    70.143 52.277    52.198
##      affectweather worried harmplants futuregen devharm harmUS personal timing
## 3631        59.441  58.563     68.802    68.367  61.792 59.132   39.911 52.296
## 3632        60.885  61.405     70.364    68.388  66.353 62.648   42.811 55.884
## 3633        52.577  57.759     63.827    63.585  60.363 57.864   42.074 52.276
## 3634        53.728  56.167     62.714    62.346  60.238 56.992   39.817 51.603
## 3635        59.384  55.738     67.269    67.289  62.891 58.861   38.786 51.005
## 3636        62.742  59.788     69.689    68.891  64.922 60.850   39.187 54.247
##         exp fundrenewables regulate CO2limits reducetax supportRPS rebates
## 3631 38.899         75.075   67.747    60.315    60.138     61.513  75.326
## 3632 43.648         75.747   68.682    64.050    62.706     61.274  75.132
## 3633 38.858         70.743   64.389    55.443    57.861     56.788  71.350
## 3634 38.495         70.961   64.135    56.283    55.571     57.649  73.226
## 3635 37.750         76.222   70.664    60.099    62.623     61.657  76.548
## 3636 40.613         78.184   71.507    63.333    62.489     62.841  77.487
##      drillANWR drilloffshore teachGW corporations president congress governor
## 3631    34.194        54.372  73.109       66.386    44.913   55.460   49.554
## 3632    30.296        46.241  74.815       68.772    50.226   58.555   47.126
## 3633    33.919        58.162  71.628       62.985    43.193   51.943   50.504
## 3634    38.233        59.662  69.438       62.077    41.326   52.308   48.237
## 3635    32.757        55.315  73.107       66.412    45.300   54.188   52.413
## 3636    35.436        54.653  72.110       70.458    50.218   60.645   54.585
```

```
##      localofficials citizens discuss mediaweekly priority
## 3631         52.229   60.608  31.819      32.064   46.924
## 3632         54.703   63.459  36.888      37.116   51.054
## 3633         52.393   58.825  29.091      31.125   47.867
## 3634         51.398   57.223  29.308      28.899   46.800
## 3635         53.331   59.631  30.378      30.213   45.687
## 3636         56.824   63.639  33.118      30.269   50.292
```

```
library(stringr)
metro[c('City', 'State')] <- str_split_fixed(metro$GeoName, ', ', 2)
chosen_vars <- c('GEOID', 'GeoName', 'TotalPop', 'happening', 'human', 'worried', 'City', 'State')
metro <- metro[, chosen_vars]
head(metro)
```

```
##      GEOID      GeoName TotalPop happening  human worried     City State
## 3631 10100 Aberdeen, SD    32710    67.750 51.553  58.563 Aberdeen    SD
## 3632 10140 Aberdeen, WA    57845    68.772 53.170  61.405 Aberdeen    WA
## 3633 10180  Abilene, TX   130201    62.129 48.285  57.759  Abilene    TX
## 3634 10220      Ada, OK    29357    65.013 49.048  56.167      Ada    OK
## 3635 10300   Adrian, MI    77534    63.537 51.129  55.738   Adrian    MI
## 3636 10420    Akron, OH   558902    70.143 52.277  59.788    Akron    OH
```

Getting voting patterns data

```
data <- read.table('president.tab', header = T, sep = "\t")
cols <- c("year", "state", "state_po", "candidate", "candidatevotes",  "totalvotes", "party_simplified")
data <- data[data$year == 2020, names(data) %in% cols]
data$party_simplified[data$party_simplified == "LIBERTARIAN"] <- "OTHER"
head(data)
```

```
##      year   state state_po           candidate candidatevotes totalvotes
## 3741 2020 ALABAMA       AL BIDEN, JOSEPH R. JR         849624    2323282
## 3742 2020 ALABAMA       AL    TRUMP, DONALD J.        1441170    2323282
## 3743 2020 ALABAMA       AL       JORGENSEN, JO          25176    2323282
## 3744 2020 ALABAMA       AL                              7312    2323282
## 3745 2020  ALASKA       AK BIDEN, JOSEPH R. JR         153778     359530
## 3746 2020  ALASKA       AK    TRUMP, DONALD J.         189951     359530
##      party_simplified
## 3741         DEMOCRAT
## 3742       REPUBLICAN
## 3743            OTHER
## 3744            OTHER
## 3745         DEMOCRAT
## 3746       REPUBLICAN
```

```
president <- data %>%
  group_by(state, state_po, party_simplified, totalvotes) %>%
  summarise(across(c(candidatevotes), sum))
```

```
## `summarise()` has grouped output by 'state', 'state_po', 'party_simplified'.
## You can override using the `.groups` argument.
```

```
president$percentage = president$candidatevotes / president$totalvotes
president
```

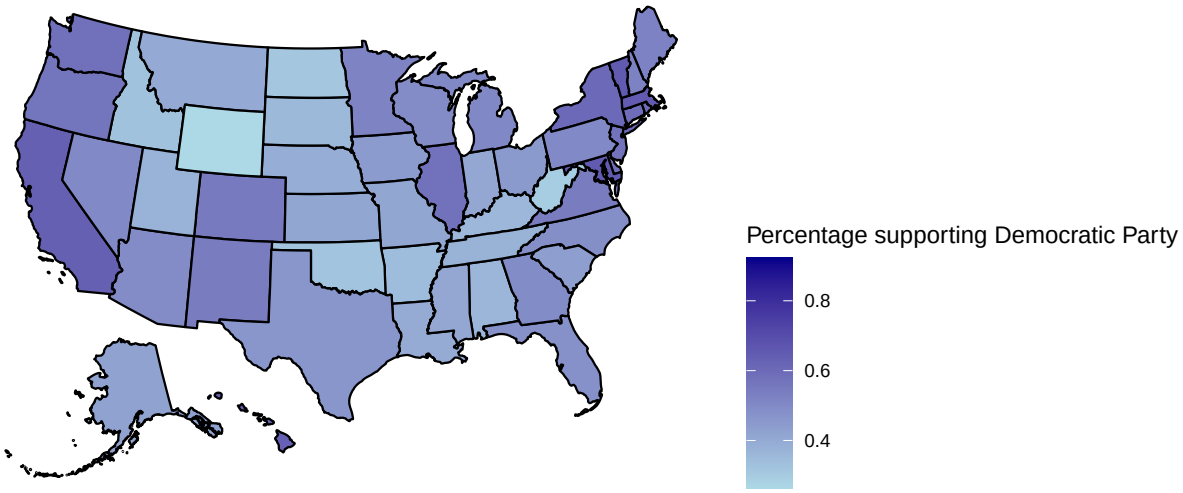
```
## # A tibble: 153 x 6
```

```
## # Groups:   state, state_po, party_simplified [153]
##    state    state_po party_simplified totalvotes candidatevotes percentage
##    <chr>    <chr>    <chr>                 <int>          <int>      <dbl>
##  1 ALABAMA  AL       DEMOCRAT            2323282         849624     0.366
##  2 ALABAMA  AL       OTHER               2323282          32488     0.0140
##  3 ALABAMA  AL       REPUBLICAN          2323282        1441170     0.620
##  4 ALASKA   AK       DEMOCRAT             359530         153778     0.428
##  5 ALASKA   AK       OTHER                359530          15801     0.0439
##  6 ALASKA   AK       REPUBLICAN           359530         189951     0.528
##  7 ARIZONA  AZ       DEMOCRAT            3387326        1672143     0.494
##  8 ARIZONA  AZ       OTHER               3387326          53497     0.0158
##  9 ARIZONA  AZ       REPUBLICAN          3387326        1661686     0.491
## 10 ARKANSAS AR       DEMOCRAT            1219069         423932     0.348
## # ... with 143 more rows
```

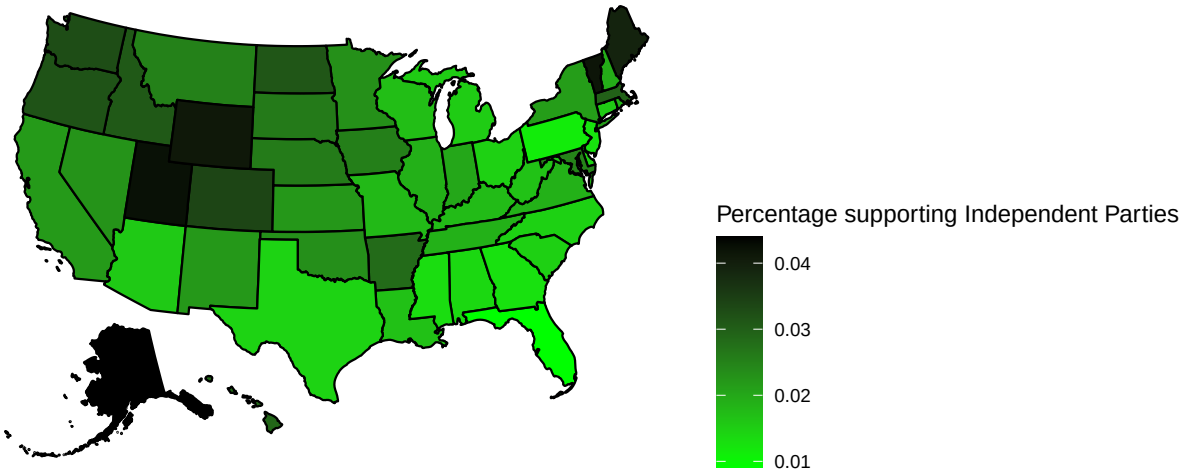
```
democrats <- president[president$party_simplified == 'DEMOCRAT', ]
head(democrats)
```

```
## # A tibble: 6 x 6
## # Groups:   state, state_po, party_simplified [6]
##   state      state_po party_simplified totalvotes candidatevotes percentage
##   <chr>      <chr>    <chr>                 <int>          <int>      <dbl>
## 1 ALABAMA    AL       DEMOCRAT            2323282         849624     0.366
## 2 ALASKA     AK       DEMOCRAT             359530         153778     0.428
## 3 ARIZONA    AZ       DEMOCRAT            3387326        1672143     0.494
## 4 ARKANSAS   AR       DEMOCRAT            1219069         423932     0.348
## 5 CALIFORNIA CA       DEMOCRAT           17500881       11110250     0.635
## 6 COLORADO   CO       DEMOCRAT            3279980        1804352     0.550
```

```
plot_usmap(data=democrats, values="percentage", color="black") +
  scale_fill_continuous(low="lightblue", high="darkblue", name='Percentage supporting Democratic Party'
  theme(legend.position = "right")
```




```
democrats <- president[president$party_simplified == 'OTHER', ]
plot_usmap(data=democrats, values="percentage", color="black") +
  scale_fill_continuous(low="green", high="black", name='Percentage supporting Independent Parties', la
  theme(legend.position = "right")
```



Merging metro table with state table (for state average) and democrat table (for democrat average)

```
merged <- merge(x=metro,y=democrats,by.x="State", by.y="state_po",all.x=TRUE)
names(merged)[names(merged) == 'State'] <- 'state_id'
head(merged)
```

```
##   state_id GEOID                  GeoName TotalPop happening  human worried
## 1       AK 27940               Juneau, AK    25286    73.590 58.538  67.986
## 2       AK 28540            Ketchikan, AK    10767    73.837 58.661  66.304
## 3       AK 21820            Fairbanks, AK    75128    70.355 53.051  63.963
## 4       AK 11260            Anchorage, AK   298632    72.739 56.138  64.196
## 5       AL 18980              Cullman, AL    64149    51.108 40.788  45.008
## 6       AL 19300 Daphne-Fairhope-Foley, AL   166595    60.503 44.677  53.962
##                    City   state party_simplified totalvotes candidatevotes
## 1                Juneau  ALASKA            OTHER     359530          15801
## 2             Ketchikan  ALASKA            OTHER     359530          15801
## 3             Fairbanks  ALASKA            OTHER     359530          15801
## 4             Anchorage  ALASKA            OTHER     359530          15801
## 5               Cullman ALABAMA            OTHER    2323282          32488
## 6 Daphne-Fairhope-Foley ALABAMA            OTHER    2323282          32488
##   percentage
## 1 0.04394904
## 2 0.04394904
## 3 0.04394904
## 4 0.04394904
## 5 0.01398367
## 6 0.01398367
```

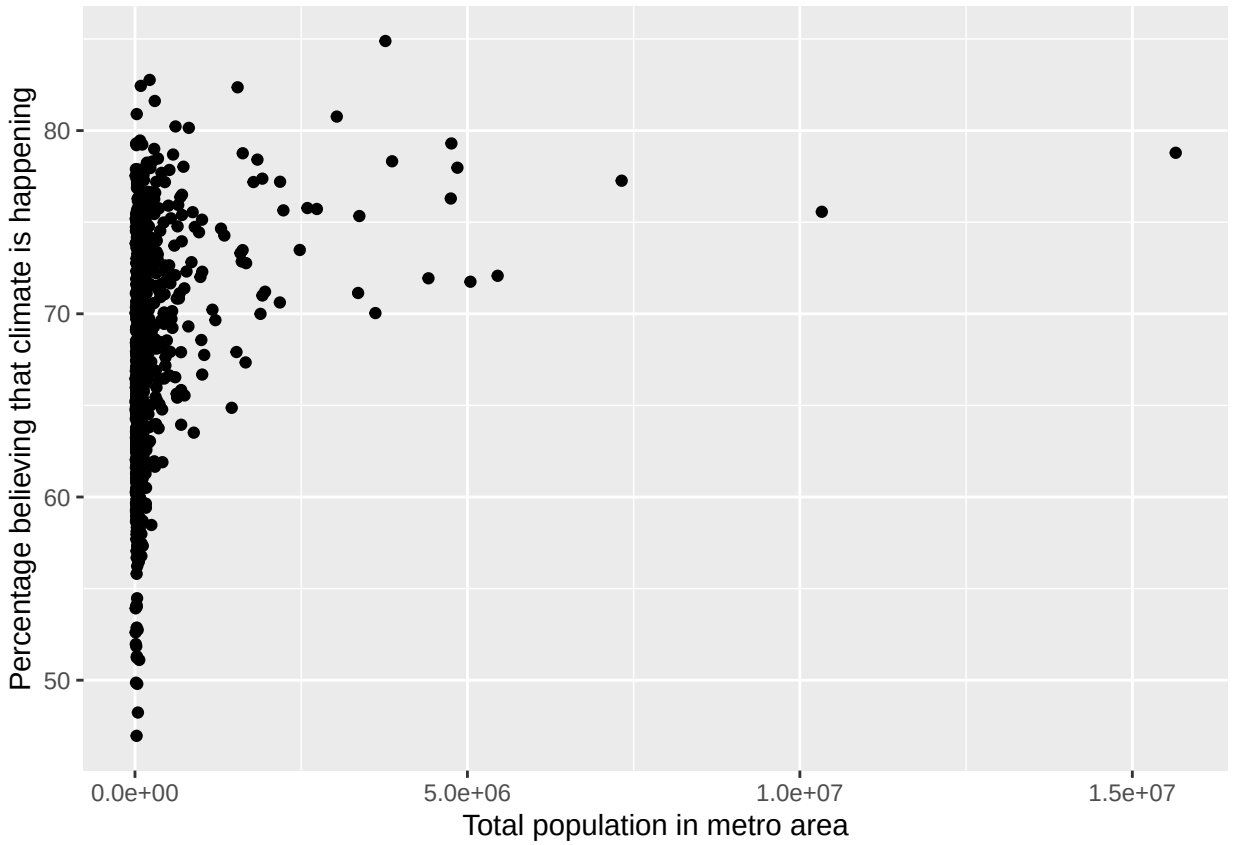
```
state$state <- toupper(state$state)
head(state)
```

```
##   GEOID      state TotalPop happening  human worried
## 2     1    ALABAMA  3779877    63.291 49.350  57.555
## 3     2     ALASKA   552673    70.188 53.277  63.234
## 4     4    ARIZONA  5414955    72.246 56.428  65.733
## 5     5   ARKANSAS  2295104    66.366 51.238  59.382
## 6     6 CALIFORNIA 30261354    77.466 63.349  72.696
## 7     8   COLORADO  4349342    72.559 57.099  65.978
```

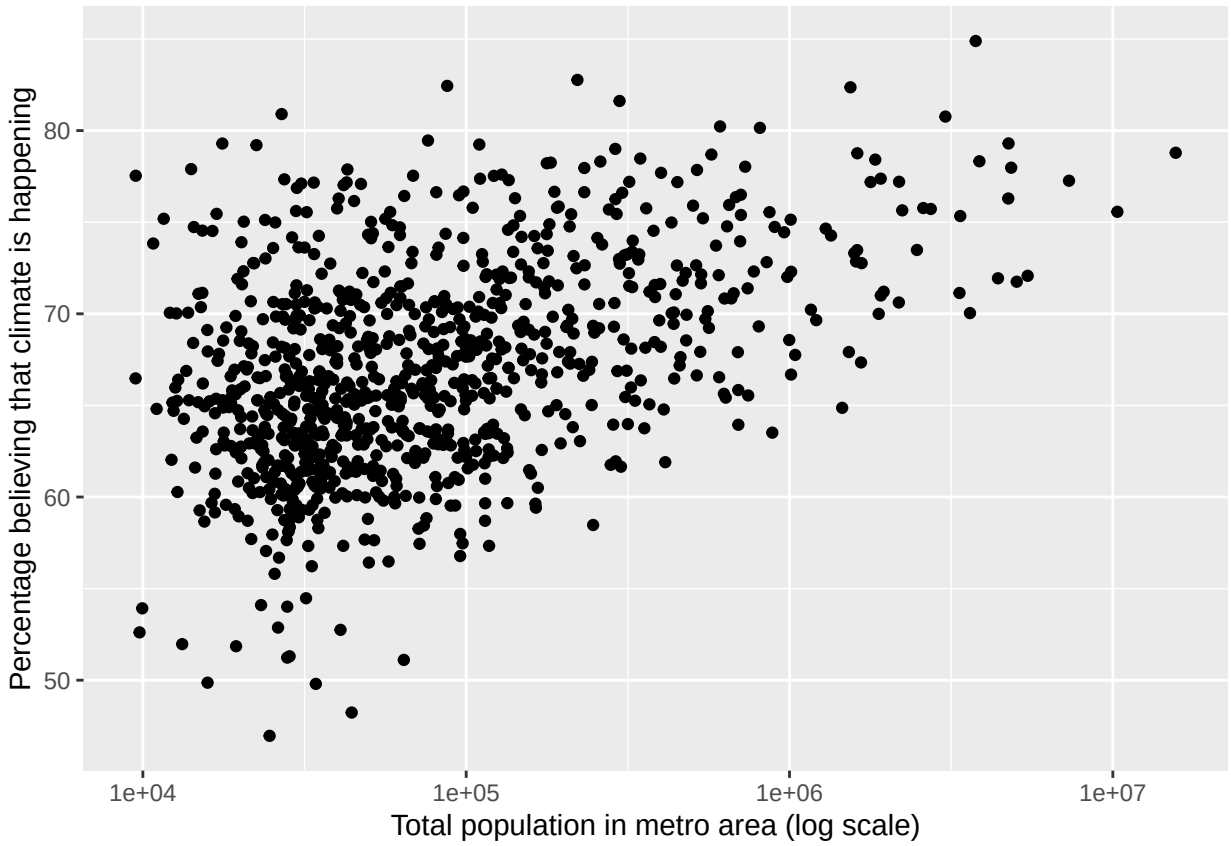
```
final <- merge(x=merged, y=state, by="state", all.x=TRUE, suffix = c("_metro", "_state"))
final$GEOID_metro <- as.factor(final$GEOID_metro)
final$GEOID_state <- as.factor(final$GEOID_state)
head(final)
```

```
##     state state_id GEOID_metro                  GeoName TotalPop_metro
## 1 ALABAMA       AL       10760        Alexander City, AL          32087
## 2 ALABAMA       AL       42460            Scottsboro, AL          40870
## 3 ALABAMA       AL       46220            Tuscaloosa, AL         191285
## 4 ALABAMA       AL       37120                 Ozark, AL          37892
## 5 ALABAMA       AL       21460            Enterprise, AL          39402
## 6 ALABAMA       AL       19300 Daphne-Fairhope-Foley, AL         166595
##   happening_metro human_metro worried_metro                 City
## 1          63.116      46.583        56.599        Alexander City
## 2          52.748      40.477        48.427            Scottsboro
## 3          66.807      49.397        58.470            Tuscaloosa
## 4          61.356      45.796        53.742                 Ozark
## 5          59.966      43.848        50.591            Enterprise
## 6          60.503      44.677        53.962 Daphne-Fairhope-Foley
##   party_simplified totalvotes candidatevotes percentage GEOID_state
## 1            OTHER    2323282          32488 0.01398367           1
## 2            OTHER    2323282          32488 0.01398367           1
## 3            OTHER    2323282          32488 0.01398367           1
## 4            OTHER    2323282          32488 0.01398367           1
## 5            OTHER    2323282          32488 0.01398367           1
## 6            OTHER    2323282          32488 0.01398367           1
##   TotalPop_state happening_state human_state worried_state
## 1        3779877          63.291       49.35        57.555
## 2        3779877          63.291       49.35        57.555
## 3        3779877          63.291       49.35        57.555
## 4        3779877          63.291       49.35        57.555
## 5        3779877          63.291       49.35        57.555
## 6        3779877          63.291       49.35        57.555
```

```
ggplot(final, aes(x=TotalPop_metro, y=happening_metro)) + geom_point() +
  labs(x = "Total population in metro area", y = "Percentage believing that climate is happening")
```

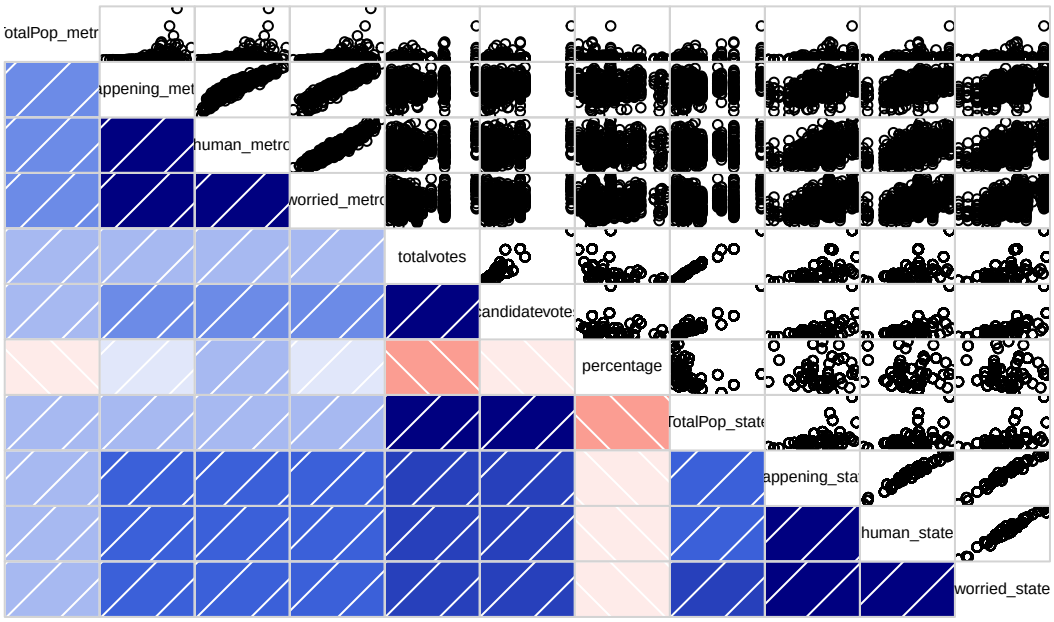



```
ggplot(final, aes(x=TotalPop_metro, y=happening_metro)) + geom_point() + scale_x_continuous(trans='log1
  labs(x = "Total population in metro area (log scale)", y = "Percentage believing that climate is happ
```




```
library(dplyr)
library(corrgram)
chosen <- final %>%
    select_if(is.numeric)
corrgram(chosen, order = NULL, lower.panel = panel.shade, upper.panel = panel.pts, text.panel = panel.t
    main = "Predictors of Climate Change Behavior Corrgram Plot")
```

Predictors of Climate Change Behavior Corrgram Plot




```
#create log column for population
chosen$population_log <- log(chosen$TotalPop_metro)
```

```
lm <- lm(happening_metro ~ population_log + percentage, data=chosen)
summary(lm)
```

```
##
## Call:
## lm(formula = happening_metro ~ population_log + percentage, data = chosen)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.4158  -3.6024  -0.2515   3.2486  15.7888
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     39.4502     1.7381  22.697  < 2e-16 ***
## population_log   2.2234     0.1444  15.401  < 2e-16 ***
## percentage     134.9187    23.7631   5.678 1.86e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.016 on 867 degrees of freedom
##   (63 observations deleted due to missingness)
## Multiple R-squared:  0.2261, Adjusted R-squared:  0.2243
## F-statistic: 126.6 on 2 and 867 DF,  p-value: < 2.2e-16
```