

Data-Driven Research

Code ▾

Hide

```
library(readr)
howe_data <- read_csv("howe-2016-data.csv")
```

Rows: 4545 Columns: 38

Column specification

Delimiter: ","

chr (2): GeoType, GeoName

dbl (36): GEOID, TotalPop, discuss, discussOppose, CO2limits, CO2limitsOppose, trustclimsciSST, trustclimsciSSTOppose, r...

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

Hide

```
head(howe_data)
```

GeoType	GE...	GeoName	TotalPop	discuss	discussOppose	CO2limits	CO2limitsOppose
<chr>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
National	9999	US	209056129	33.495	31.081	69.376	29.386
State	1	Alabama	3217902	27.468	38.892	63.122	35.589
State	2	Alaska	460319	36.632	28.074	68.231	31.123
State	4	Arizona	4284776	36.724	28.882	66.001	32.625
State	5	Arkansas	1949963	28.832	37.666	63.744	35.166
State	6	California	24865866	39.309	25.920	73.740	24.959

6 rows | 1-8 of 38 columns

Hide

```
unique(howe_data$GeoType)
```

```
[1] "National" "State"    "County"   "cd113"    "CBSA"
```

Hide

```
# national: US (1)
# state (51)
# county (3142)
# cd113: congressional district (435)
# CBSA: Core-based statistical area (916)
```

Hide

```
# filter the data regarding scales
national <- howe_data[howe_data$GeoType=='National',]
county <- howe_data[howe_data$GeoType=='County',]
cd <- howe_data[howe_data$GeoType=='cd113',]
state <- howe_data[howe_data$GeoType=='State',]
cbsa <- howe_data[howe_data$GeoType=='CBSA',]
```

Hide

```
# groups column names into vectors
pol_pref <- c('CO2limits', 'regulate', 'supportRPS', 'fundrenewables')
pol_pref_o <- c('CO2limitsOppose', 'regulateOppose', 'supportRPSOppose', 'fundrenewablesOppose')
belief <- c('happening', 'human', 'consensus')
belief_o <- c('happeningOppose', 'humanOppose', 'consensusOppose', 'trustclimsciSStopose')
risk <- c('worried', 'personal', 'harmUS', 'devharm', 'futuregen', 'timing', 'harmlplants')
risk_o <- c('worriedOppose', 'personalOppose', 'harmUSOppose', 'devharmOppose', 'futuregenOppose', 'timingOppose', 'harmlplantsOppose')
behavior <- c('discuss', 'mediaweekly')
behavior_o <- c('discussOppose', 'mediaweeklyOppose')
pol_bel <- c(pol_pref, belief)

all <- c(pol_pref, belief, risk, behavior)
all_o <- c(pol_pref_o, belief_o, risk_o, behavior_o)
```

Hide

```
library(dplyr)
library(corrgram)

# mutate county level data to include different columns

county_all <-
county %>%
  select(all_of(all))

county_all_o <-
county %>%
  select(all_of(all_o))

county_p <-
county %>%
  select(all_of(c(pol_pref, pol_pref_o)))

county_f <-
county %>%
  select(all_of(c('discuss', 'happening', 'human', 'consensus')))
```

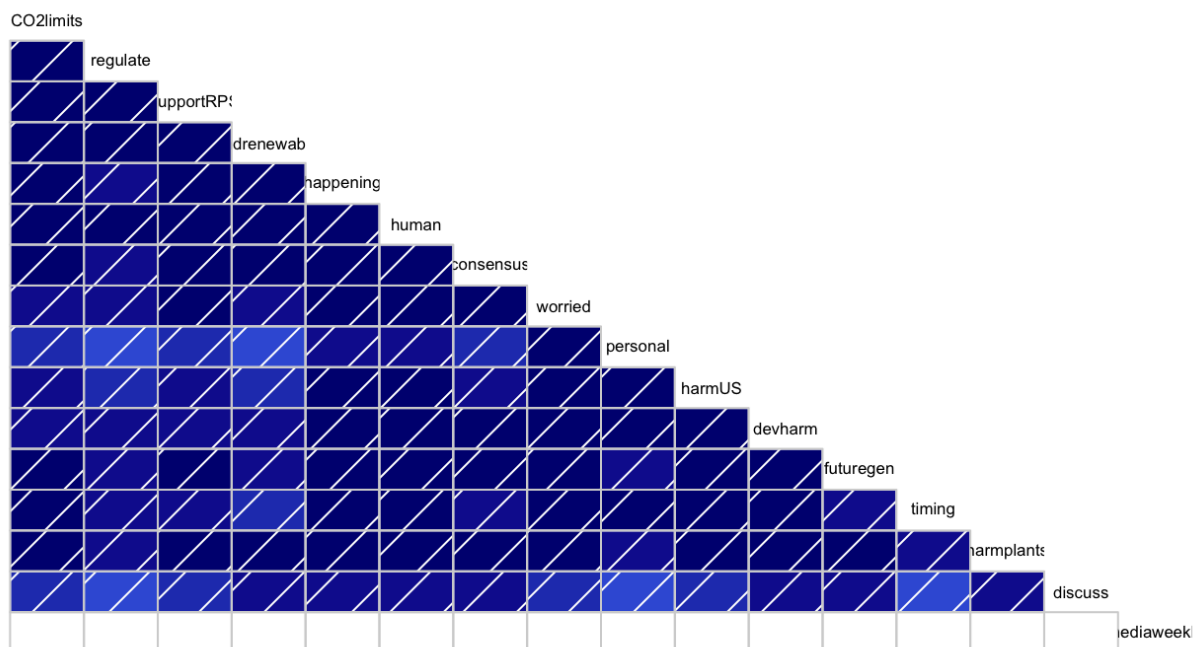
[Hide](#)

```
# explore correlation with all columns in county
all_county = cor(county_all)
round(all_county, 2)
```

	CO2limits	regulate	supportRPS	fundrenewables	happening	human	consensus	worried	personal	harmUS	devharm
CO2limits	1.00	0.90	0.91	0.86	0.93	0.91	0.91	0.85	0.69	0.80	0.84
regulate	0.90	1.00	0.95	0.89	0.85	0.87	0.84	0.77	0.53	0.68	0.73
supportRPS	0.91	0.95	1.00	0.89	0.93	0.95	0.91	0.86	0.67	0.80	0.85
fundrenewables	0.86	0.89	0.89	1.00	0.87	0.86	0.92	0.74	0.46	0.68	0.77
happening	0.93	0.85	0.93	0.87	1.00	0.97	0.95	0.94	0.80	0.92	0.94
human	0.91	0.87	0.95	0.86	0.97	1.00	0.94	0.94	0.79	0.91	0.94
consensus	0.91	0.84	0.91	0.92	0.95	0.94	1.00	0.86	0.66	0.84	0.90
worried	0.85	0.77	0.86	0.74	0.94	0.94	0.86	1.00	0.90	0.98	0.97
personal	0.69	0.53	0.67	0.46	0.80	0.79	0.66	0.90	1.00	0.94	0.87
harmUS	0.80	0.68	0.80	0.68	0.92	0.91	0.84	0.98	0.94	1.00	0.97
devharm	0.84	0.73	0.85	0.77	0.94	0.94	0.90	0.97	0.87	0.97	1.00
futuregen	0.86	0.78	0.87	0.85	0.95	0.95	0.92	0.94	0.76	0.91	0.97
timing	0.86	0.75	0.82	0.66	0.91	0.87	0.80	0.93	0.92	0.93	0.88
harmplants	0.89	0.80	0.87	0.86	0.95	0.94	0.92	0.93	0.76	0.90	0.96
discuss	0.64	0.49	0.60	0.72	0.75	0.73	0.84	0.70	0.54	0.71	0.78
mediaweekly	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	futuregen	timing	harmplants	discuss	mediaweekly						
CO2limits	0.86	0.86	0.89	0.64	NA						
regulate	0.78	0.75	0.80	0.49	NA						
supportRPS	0.87	0.82	0.87	0.60	NA						
fundrenewables	0.85	0.66	0.86	0.72	NA						
happening	0.95	0.91	0.95	0.75	NA						
human	0.95	0.87	0.94	0.73	NA						
consensus	0.92	0.80	0.92	0.84	NA						
worried	0.94	0.93	0.93	0.70	NA						
personal	0.76	0.92	0.76	0.54	NA						
harmUS	0.91	0.93	0.90	0.71	NA						
devharm	0.97	0.88	0.96	0.78	NA						
futuregen	1.00	0.82	0.99	0.79	NA						
timing	0.82	1.00	0.84	0.57	NA						
harmplants	0.99	0.84	1.00	0.78	NA						
discuss	0.79	0.57	0.78	1.00	NA						
mediaweekly	NA	NA	NA	NA	1						

Hide

```
corrgram(all_county, order = NULL, lower.panel = panel.shade, upper.panel = NULL)
```



Hide

```
state_all <-
state %>%
  select(all_of(all))

state_all_o <-
state %>%
  select(all_of(all_o))
```

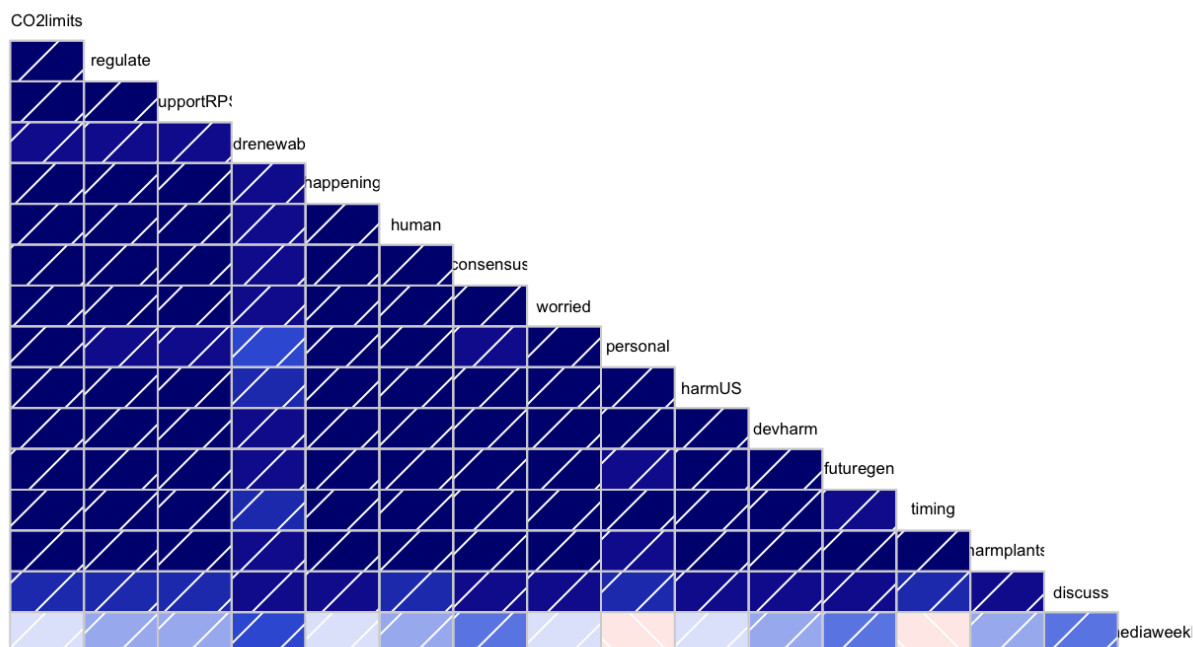
Hide

```
# explore correlation in all state data
all_state = cor(state_all)
round(all_state, 2)
```

	CO2limits	regulate	supportRPS	fundrenewables	happening	human	consensus	worried	personal	harmUS	devharm	futuregen	timing	harmplants	discuss	mediaweekly
CO2limits	1.00	0.96	0.95	0.74	0.97	0.95	0.92	0.97	0.88	0.94	0.94	0.93	0.91	0.95	0.70	0.03
regulate	0.97	1.00	0.97	0.84	0.96	0.97	0.97	0.95	0.93	0.95	0.95	0.93	0.88	0.96	0.70	0.19
supportRPS	0.95	0.83	1.00	0.79	0.96	0.98	0.97	0.95	0.85	0.94	0.96	0.97	0.90	0.95	0.71	0.23
fundrenewables	0.74	0.52	0.79	1.00	0.79	0.77	0.83	0.74	0.52	0.66	0.76	0.82	0.58	0.83	0.72	0.47
happening	0.97	0.96	0.96	0.79	1.00	0.96	0.97	0.99	0.92	0.97	0.98	0.95	0.94	0.98	0.80	0.13
human	0.95	0.97	0.98	0.77	0.96	1.00	0.96	0.97	0.89	0.96	0.96	0.98	0.98	0.95	0.70	0.18
consensus	0.92	0.97	0.97	0.83	0.97	0.96	1.00	0.95	0.85	0.95	0.98	0.95	0.92	0.98	0.81	0.32
worried	0.97	0.95	0.95	0.74	0.99	0.97	0.95	1.00	0.93	0.98	0.97	0.95	0.94	0.96	0.78	0.09
personal	0.88	0.83	0.85	0.52	0.92	0.89	0.85	0.93	1.00	0.97	0.90	0.82	0.98	0.84	0.66	-0.11
harmUS	0.94	0.93	0.94	0.66	0.97	0.96	0.95	0.98	0.97	1.00	0.97	0.97	0.98	0.93	0.75	0.06
devharm	0.94	0.95	0.96	0.76	0.98	0.96	0.98	0.97	0.90	0.97	1.00	0.97	0.91	0.98	0.81	0.22
futuregen	0.93	0.97	0.97	0.82	0.95	0.97	0.98	0.95	0.82	0.93	0.97	1.00	0.85	0.98	0.77	0.34
timing	0.91	0.88	0.90	0.58	0.94	0.92	0.89	0.94	0.98	0.98	0.91	0.85	1.00	0.87	0.67	-0.09
harmplants	0.95	0.96	0.95	0.83	0.98	0.95	0.98	0.96	0.84	0.93	0.98	0.98	0.87	1.00	0.82	0.27
discuss	0.70	0.70	0.71	0.72	0.80	0.70	0.81	0.78	0.66	0.75	0.81	0.77	0.67	0.82	1.00	0.41
mediaweekly	0.03	0.19	0.23	0.47	0.13	0.18	0.32	0.09	-0.11	0.06	0.22	0.34	-0.09	0.27	0.41	1.00
	futuregen	timing	harmplants	discuss	mediaweekly											
CO2limits	0.93	0.91	0.95	0.70	0.03											
regulate	0.97	0.88	0.96	0.70	0.19											
supportRPS	0.97	0.90	0.95	0.71	0.23											
fundrenewables	0.82	0.58	0.83	0.72	0.47											
happening	0.95	0.94	0.98	0.80	0.13											
human	0.97	0.92	0.95	0.70	0.18											
consensus	0.98	0.89	0.98	0.81	0.32											
worried	0.95	0.94	0.96	0.78	0.09											
personal	0.82	0.98	0.84	0.66	-0.11											
harmUS	0.93	0.98	0.93	0.75	0.06											
devharm	0.97	0.91	0.98	0.81	0.22											
futuregen	1.00	0.85	0.98	0.77	0.34											
timing	0.85	1.00	0.87	0.67	-0.09											
harmplants	0.98	0.87	1.00	0.82	0.27											
discuss	0.77	0.67	0.82	1.00	0.41											
mediaweekly	0.34	-0.09	0.27	0.41	1.00											

Hide

```
corrgram(all_state, order = NULL, lower.panel = panel.shade, upper.panel = NULL)
```



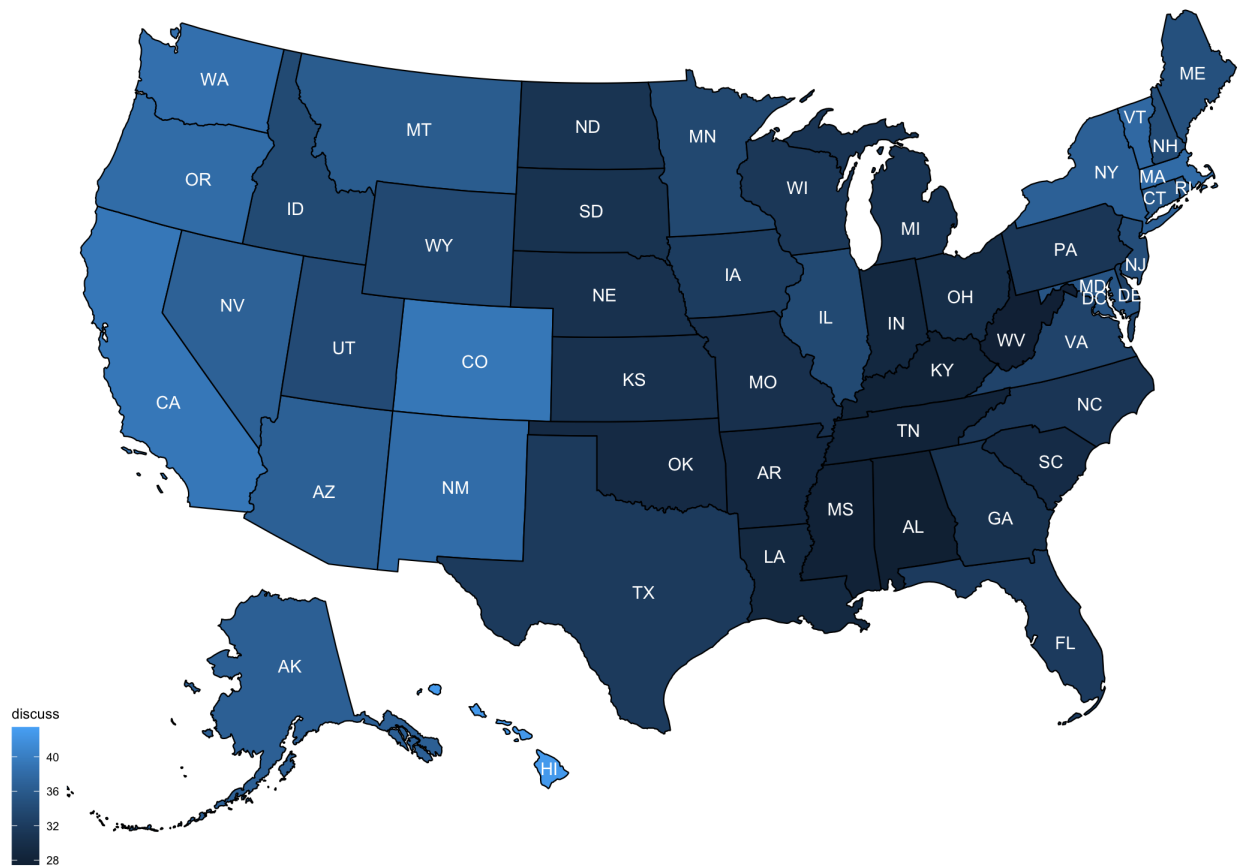
Hide

```
# rename the column to match the usmap library
names(state)[names(state) == 'GeoName'] <- 'state'
```

Hide

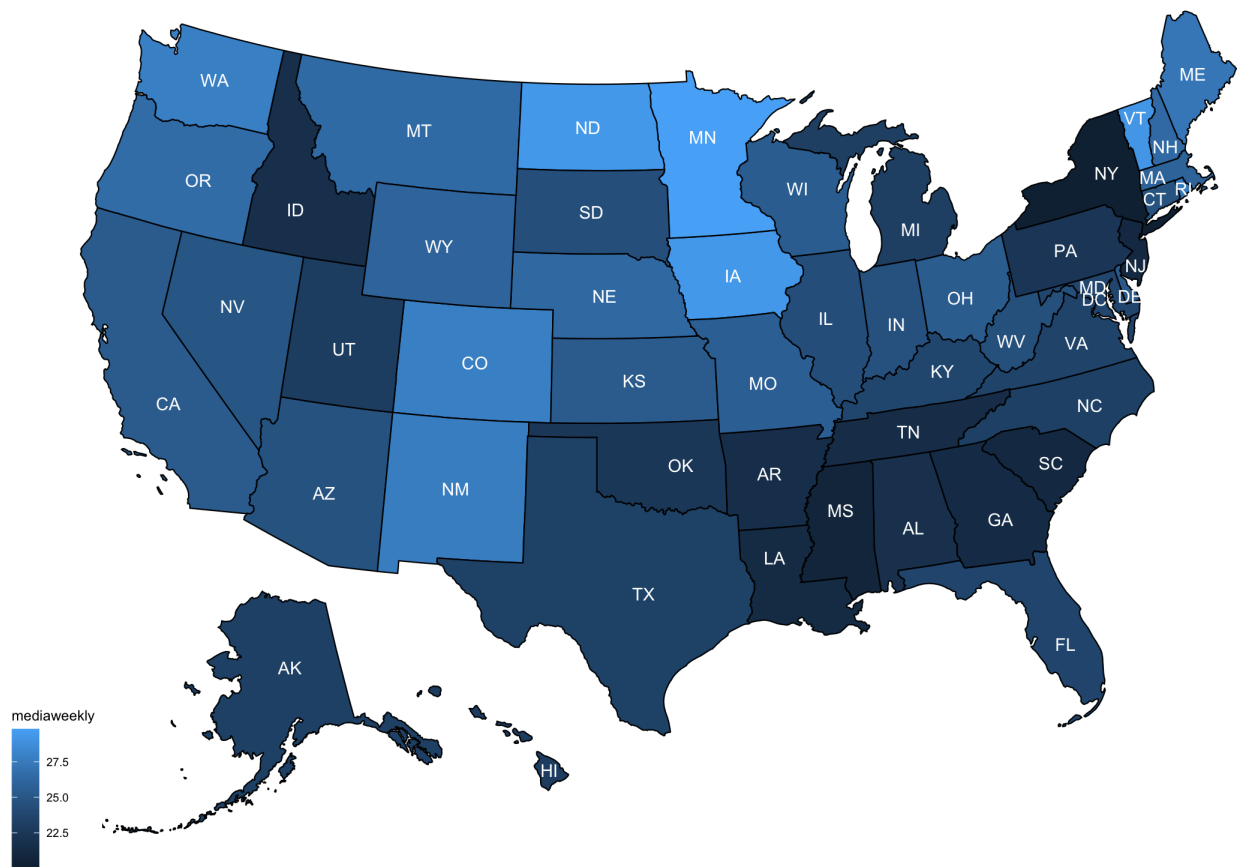
```
library(usmap)

# Color maps with data
plot_usmap(data = state, values = "discuss", labels = TRUE, label_color = "white")
```



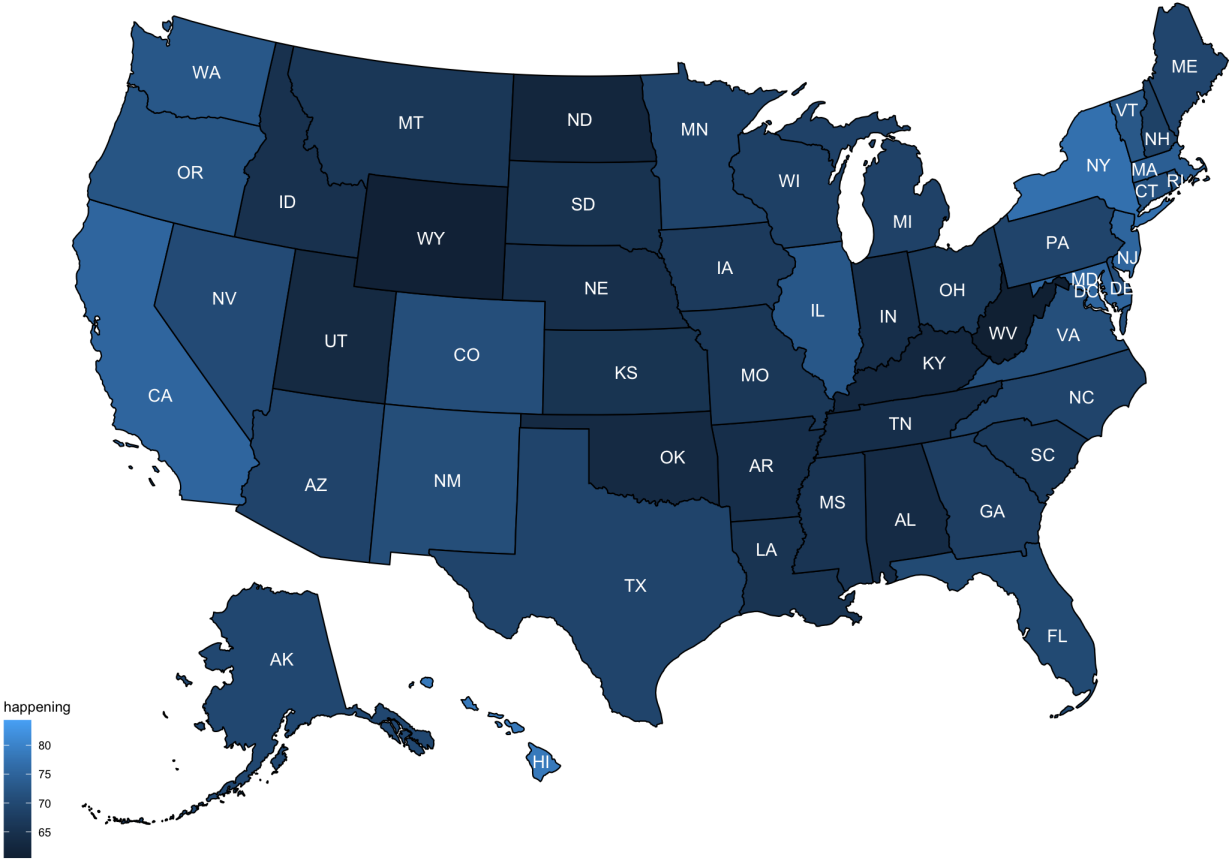
Hide

```
plot_usmap(data = state, values = "mediaweekly", labels = TRUE, label_color = "white" )
```



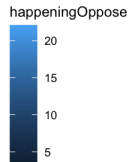
Hide

```
plot_usmap(data = state, values = "happening", labels = TRUE, label_color = "white")
```



Hide

```
plot_usmap(data = state, values = "happeningOppose", labels = TRUE, label_color = "white")
```



```
# plot_usmap(data = state, values = "worried", labels = TRUE, label_color = "white")
# plot_usmap(data = state, values = "trustclimsciSST", labels = TRUE, label_color =
  "white")
```

```
library(forcats)
library(reshape2)
```

10/13

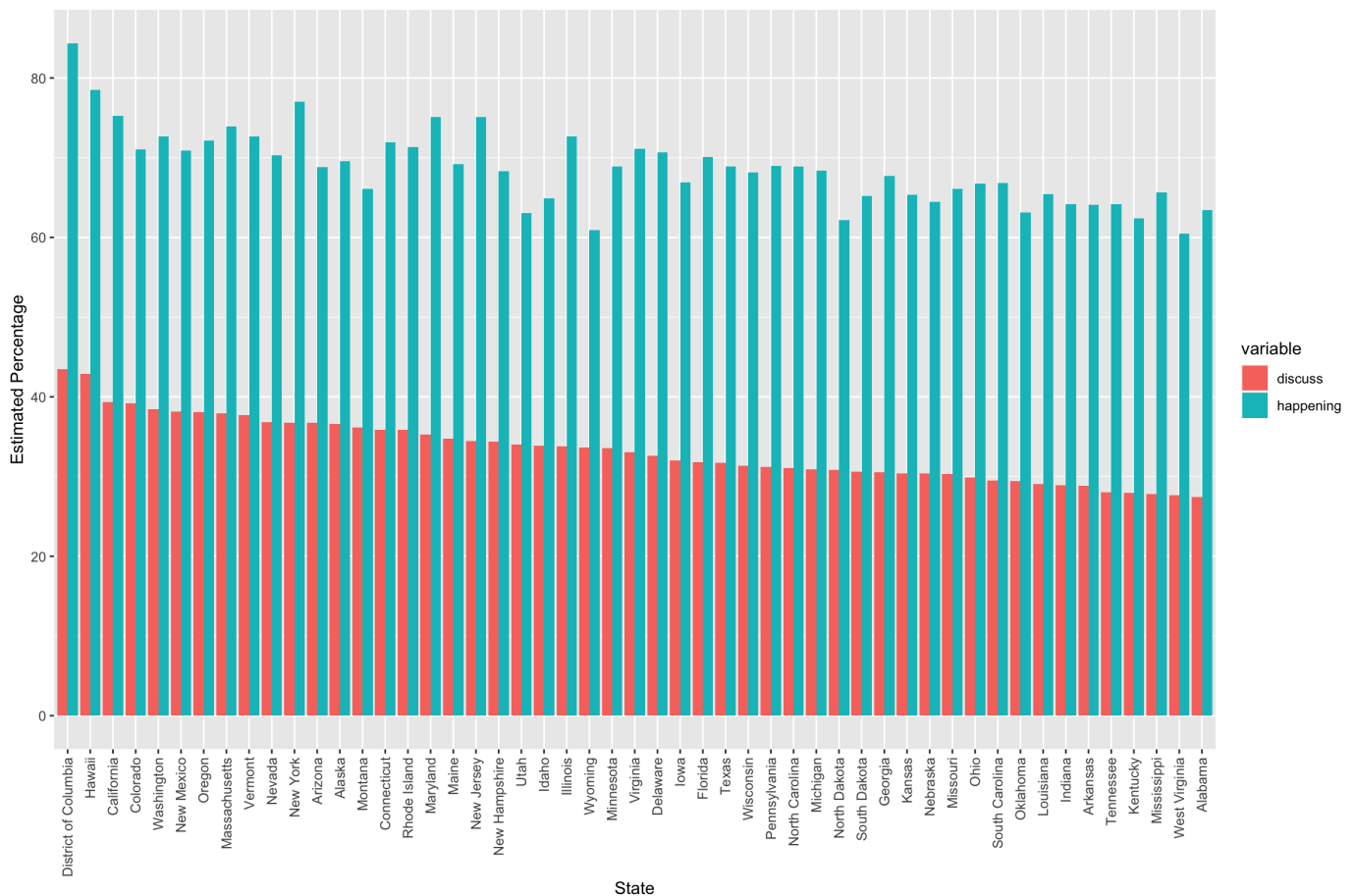
```
# plot bar chart with descending discuss
state_s <-
state %>%
  select(all_of(c('state', 'discuss', 'mediaweekly', 'happening')))

state_s <-
state_s %>%
  mutate(state = fct_reorder(state, desc(discuss)))

# load the library
library(forcats)

# Reorder following the value of another column:
statec <- melt(state_s[,c('state', 'discuss', 'happening')],id.vars = 1)

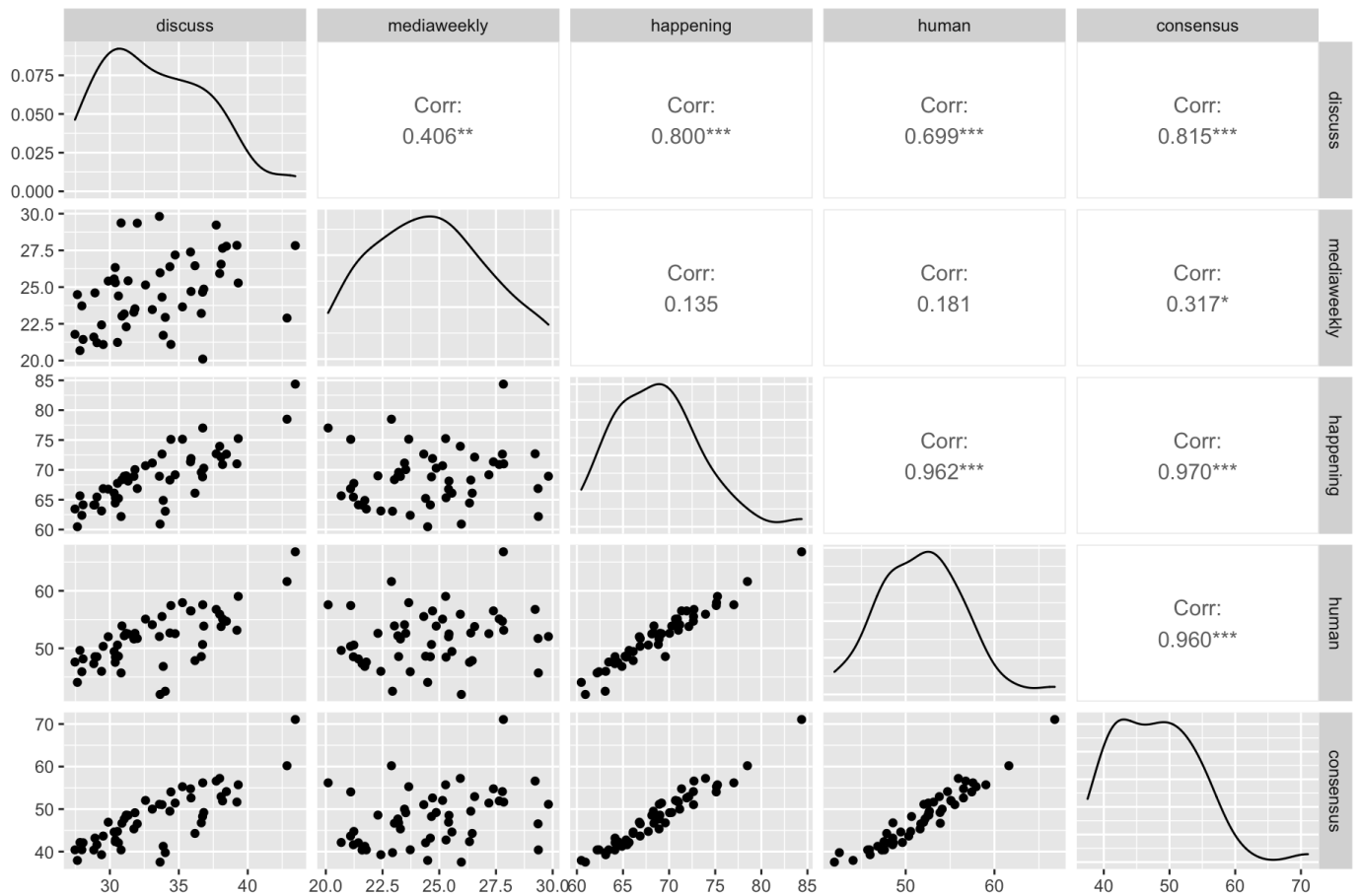
ggplot(statec,aes(x = state,y = value)) +
  geom_bar(aes(fill = variable),stat = "identity",position = "dodge") + theme(axes.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) + xlab("State") + ylab("Estimated Percentage")
```


[Hide](#)

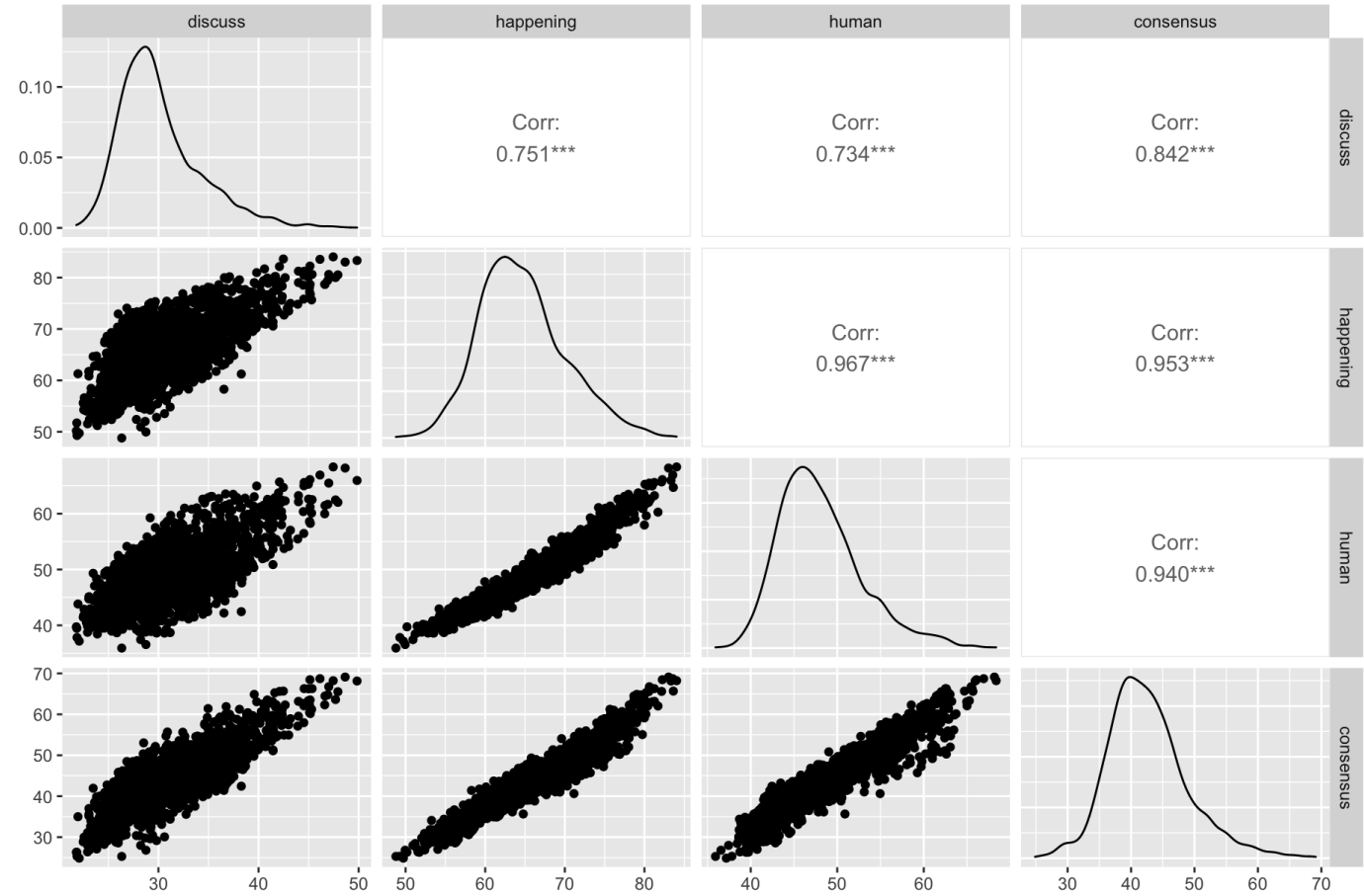
```
# Quick display of two capabilities of GGally, to assess the distribution and correlation of variables
library(GGally)
```

```
Registered S3 method overwritten by 'GGally':
  method from
+ .gg      ggplot2
```

```
ggpairs(  
  data = state_f,  
  lower = list(continuous = "points", combo = "dot_no_facet")  
)
```



```
ggpairs(  
  data = county_f,  
  lower = list(continuous = "points", combo = "dot_no_facet")  
)
```



...