G-Trends Analysis

Chaoqun Yin

2/16/2019

Prepare the data

```
gm <- read_csv("Data/geoMap.csv", skip = 2)
```

```
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `gift for boyfriend: (2/14/13 - 2/14/18)` = col_integer(),
##   `gift for girlfriend: (2/14/13 - 2/14/18)` = col_character()
## )
```

```
colnames(gm) <- c("Region", "GB", "GG")
gm$GB <- as.numeric(gm$GB)
gm$GG <- as.numeric(gm$GG)
```

```
## Warning: NAs introduced by coercion
```

Q1: Which are the states where GG is smaller than 1? Find those and replace them with zero.

```
gm <- replace(gm, is.na(gm), 0)
gm$Region[which(gm$GG == 0)]
```

```
## [1] "South Dakota" "Maine"        "Idaho"
```

Q2: For How Many States GB > GG?

```
attach(gm)
n = 0
for (i in 1:length(Region)){
  if (GB[i] > GG[i]){
    n = n + 1
  }
}
print(n)
```

```
## [1] 46
```

Q3: Find any states where GG+10 > GB

```
n = 0
for (i in 1:length(gm$Region)){
  if (gm$GB[i] < as.numeric(gm$GG[i]) + 10){
```

```r
    n = n + 1
    print(gm$Region[i])
  }
}
```

```
## [1] "Washington"
```

Q4: What is the % of states for which GG+10 > GB?

```r
n = 0
for (i in 1:length(gm$Region)){
  if (gm$GB[i] < as.numeric(gm$GG[i]) + 10){
    n = n + 1
  }
}
cat((n/length(gm$Region))*100, "%")
```

```
## 2.173913 %
```

Q5: What is the ratio GG/GB for the state of New Hampshire?

```r
i = length(which(gm$Region == "New Hampshire"))
print(gm$GG[i]/gm$GB[i])
```

```
## [1] 0.5
```

Q6: Create a Bar Plot of GG & GB values for each state.

```r
library(ggplot2)
library(reshape2)
gm_bar <- melt(gm, id.vars = "Region")
head(gm_bar)
```

```
##          Region variable value
## 1 New Hampshire       GB   100
## 2      New York       GB    99
## 3  South Dakota       GB    88
## 4        Hawaii       GB    87
## 5  Pennsylvania       GB    85
## 6      Michigan       GB    84
```

```r
ggplot(gm_bar, aes(x = Region, y = value, fill = variable)) +
  geom_bar(stat='identity', position='dodge') + coord_flip()
```