DSCI353-353m-453: Class 14a Mapping Google Trends data For Thanksgiving

2008-351-351m-451-W14a-p2-GIS-Gtrends-Thnksgvng-Maps

Roger H. French, Raymond Weiser, Sameera Nalin Venkat

29 November, 2022

Contents

14.1.4 Mapping Google Trends data For Thanksgiving

Lets use the R package `usmap`

- which enables incredibly easy and fast
  - creation of US maps in R.

In honor of US Thanksgiving,

- we'll use the `gtrendsR` package
  - to pull US Google search results
    * on the keyword "thanksgiving"
  - and plot the popularity by state.

```
library(gtrendsR)
library(tidyverse)
```

14.1.4.1 Install and Load Packages

```
## -- Attaching packages ---------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2

## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(cdlTools)
library(usmap)
```

```
##
## Attaching package: 'usmap'

## The following object is masked from 'package:cdlTools':
##
##     fips
```
```r
#Set a color
orange <- "#C9592E"
```

14.1.4.2 Use Google Trends to get thanksgiving trends

- Use the gTrendsR Package

    - to get the query trends for thanksgiving in the US
    - for the past 24 hours

```r
thanksgiving <-
  gtrendsR::gtrends(keyword = "thanksgiving", geo = "US", time = "now 1-d") # last day

glimpse(thanksgiving)
```

```
## List of 7
##  $ interest_over_time :'data.frame': 179 obs. of  7 variables:
##   ..$ date    : POSIXct[1:179], format: "2022-11-28 14:56:00" "2022-11-28 15:04:00" ...
##   ..$ hits    : int [1:179] 100 98 97 94 99 89 95 89 85 87 ...
##   ..$ keyword : chr [1:179] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:179] "US" "US" "US" "US" ...
##   ..$ time    : chr [1:179] "now 1-d" "now 1-d" "now 1-d" "now 1-d" ...
##   ..$ gprop   : chr [1:179] "web" "web" "web" "web" ...
##   ..$ category: int [1:179] 0 0 0 0 0 0 0 0 0 0 ...
##  $ interest_by_country: NULL
##  $ interest_by_region :'data.frame': 51 obs. of  5 variables:
##   ..$ location: chr [1:51] "Utah" "California" "Pennsylvania" "New Jersey" ...
##   ..$ hits    : int [1:51] 100 73 73 72 71 71 71 71 70 69 ...
##   ..$ keyword : chr [1:51] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:51] "US" "US" "US" "US" ...
##   ..$ gprop   : chr [1:51] "web" "web" "web" "web" ...
##  $ interest_by_dma    :'data.frame': 210 obs. of  5 variables:
##   ..$ location: chr [1:210] "Alpena MI" "Salt Lake City UT" "Sherman TX-Ada OK" "St. Joseph MO" ...
##   ..$ hits    : int [1:210] 100 71 68 63 58 57 56 56 55 54 ...
##   ..$ keyword : chr [1:210] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:210] "US" "US" "US" "US" ...
##   ..$ gprop   : chr [1:210] "web" "web" "web" "web" ...
##  $ interest_by_city   :'data.frame': 200 obs. of  5 variables:
##   ..$ location: chr [1:200] "Pottsboro" "Highland" "Maltby" "Fayetteville" ...
##   ..$ hits    : int [1:200] NA NA NA NA 100 NA 75 NA NA NA ...
##   ..$ keyword : chr [1:200] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:200] "US" "US" "US" "US" ...
##   ..$ gprop   : chr [1:200] "web" "web" "web" "web" ...
##  $ related_topics     :'data.frame': 38 obs. of  6 variables:
##   ..$ subject       : chr [1:38] "100" "9" "8" "6" ...
##   ..$ related_topics: chr [1:38] "top" "top" "top" "top" ...
##   ..$ value         : chr [1:38] "Thanksgiving" "Thanksgiving dinner" "Day" "Turkey meat" ...
##   ..$ geo           : chr [1:38] "US" "US" "US" "US" ...
##   ..$ keyword       : chr [1:38] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
```

```
##    ..$ category       : int [1:38] 0 0 0 0 0 0 0 0 0 0 ...
##    ..- attr(*, "reshapeLong")=List of 4
##    .. ..$ varying:List of 1
##    .. .. ..- attr(*, "v.names")= chr "value"
##    .. .. ..- attr(*, "times")= chr "top"
##    .. ..$ v.names: chr "value"
##    .. ..$ idvar  : chr "id"
##    .. ..$ timevar: chr "related_topics"
##  $ related_queries   :'data.frame': 50 obs. of  6 variables:
##    ..$ subject       : chr [1:50] "100" "88" "58" "57" ...
##    ..$ related_topics: chr [1:50] "top" "top" "top" "top" ...
##    ..$ value         : chr [1:50] "thanksgiving 2022" "thanksgiving day" "turkey thanksgiving" "turkey
##    ..$ geo           : chr [1:50] "US" "US" "US" "US" ...
##    ..$ keyword       : chr [1:50] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ category      : int [1:50] 0 0 0 0 0 0 0 0 0 0 ...
##    ..- attr(*, "reshapeLong")=List of 4
##    .. ..$ varying:List of 1
##    .. .. ..- attr(*, "v.names")= chr "value"
##    .. .. ..- attr(*, "times")= chr "top"
##    .. ..$ v.names: chr "value"
##    .. ..$ idvar  : chr "id"
##    .. ..$ timevar: chr "related_topics"
##  - attr(*, "class")= chr [1:2] "gtrends" "list"
```

```r
thanksgiving_2021 <-
  gtrends(keyword = "thanksgiving", geo = "US", time = "2021-11-23 2021-11-25", tz = 0 ) # last year

glimpse(thanksgiving_2021)
```

```
## List of 7
##  $ interest_over_time :'data.frame': 3 obs. of  7 variables:
##    ..$ date    : POSIXct[1:3], format: "2021-11-23" "2021-11-24" ...
##    ..$ hits    : int [1:3] 14 25 100
##    ..$ keyword : chr [1:3] "thanksgiving" "thanksgiving" "thanksgiving"
##    ..$ geo     : chr [1:3] "US" "US" "US"
##    ..$ time    : chr [1:3] "2021-11-23 2021-11-25" "2021-11-23 2021-11-25" "2021-11-23 2021-11-25"
##    ..$ gprop   : chr [1:3] "web" "web" "web"
##    ..$ category: int [1:3] 0 0 0
##  $ interest_by_country: NULL
##  $ interest_by_region :'data.frame': 51 obs. of  5 variables:
##    ..$ location: chr [1:51] "Florida" "North Carolina" "New Hampshire" "Massachusetts" ...
##    ..$ hits    : int [1:51] 100 95 93 90 90 90 89 89 88 88 ...
##    ..$ keyword : chr [1:51] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ geo     : chr [1:51] "US" "US" "US" "US" ...
##    ..$ gprop   : chr [1:51] "web" "web" "web" "web" ...
##  $ interest_by_dma    :'data.frame': 210 obs. of  5 variables:
##    ..$ location: chr [1:210] "Orlando-Daytona Beach-Melbourne FL" "Miami-Ft. Lauderdale FL" "Panama C
##    ..$ hits    : int [1:210] 100 97 94 92 90 90 90 89 89 88 ...
##    ..$ keyword : chr [1:210] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ geo     : chr [1:210] "US" "US" "US" "US" ...
##    ..$ gprop   : chr [1:210] "web" "web" "web" "web" ...
##  $ interest_by_city   :'data.frame': 200 obs. of  5 variables:
##    ..$ location: chr [1:200] "Sylvania" "Stockton" "Saint Simons Island" "Nicholasville" ...
##    ..$ hits    : int [1:200] NA NA NA NA NA NA NA NA NA NA ...
##    ..$ keyword : chr [1:200] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
```

```
##    ..$ geo     : chr [1:200] "US" "US" "US" "US" ...
##    ..$ gprop   : chr [1:200] "web" "web" "web" "web" ...
##  $ related_topics    :'data.frame': 20 obs. of  6 variables:
##    ..$ subject      : chr [1:20] "100" "14" "7" "6" ...
##    ..$ related_topics: chr [1:20] "top" "top" "top" "top" ...
##    ..$ value        : chr [1:20] "Thanksgiving" "History" "Macy's Thanksgiving Day Parade" "Thanksgi
##    ..$ geo          : chr [1:20] "US" "US" "US" "US" ...
##    ..$ keyword      : chr [1:20] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ category     : int [1:20] 0 0 0 0 0 0 0 0 0 0 ...
##    ..- attr(*, "reshapeLong")=List of 4
##    .. ..$ varying:List of 1
##    .. .. ..- attr(*, "v.names")= chr "value"
##    .. .. ..- attr(*, "times")= chr "top"
##    .. ..$ v.names: chr "value"
##    .. ..$ idvar  : chr "id"
##    .. ..$ timevar: chr "related_topics"
##  $ related_queries   :'data.frame': 50 obs. of  6 variables:
##    ..$ subject      : chr [1:50] "100" "97" "87" "86" ...
##    ..$ related_topics: chr [1:50] "top" "top" "top" "top" ...
##    ..$ value        : chr [1:50] "history of thanksgiving" "thanksgiving history" "happy thanksgiving
##    ..$ geo          : chr [1:50] "US" "US" "US" "US" ...
##    ..$ keyword      : chr [1:50] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ category     : int [1:50] 0 0 0 0 0 0 0 0 0 0 ...
##    ..- attr(*, "reshapeLong")=List of 4
##    .. ..$ varying:List of 1
##    .. .. ..- attr(*, "v.names")= chr "value"
##    .. .. ..- attr(*, "times")= chr "top"
##    .. ..$ v.names: chr "value"
##    .. ..$ idvar  : chr "id"
##    .. ..$ timevar: chr "related_topics"
##  - attr(*, "class")= chr [1:2] "gtrends" "list"
```

```r
thanksgiving_2020 <-
  gtrends(keyword = "thanksgiving", geo = "US", time = "2020-11-26 2020-11-27") # 2nd year

glimpse(thanksgiving_2020)
```

```
## List of 7
##  $ interest_over_time :'data.frame': 2 obs. of  7 variables:
##    ..$ date    : POSIXct[1:2], format: "2020-11-26" "2020-11-27"
##    ..$ hits    : int [1:2] 100 21
##    ..$ keyword : chr [1:2] "thanksgiving" "thanksgiving"
##    ..$ geo     : chr [1:2] "US" "US"
##    ..$ time    : chr [1:2] "2020-11-26 2020-11-27" "2020-11-26 2020-11-27"
##    ..$ gprop   : chr [1:2] "web" "web"
##    ..$ category: int [1:2] 0 0
##  $ interest_by_country: NULL
##  $ interest_by_region :'data.frame': 51 obs. of  5 variables:
##    ..$ location: chr [1:51] "Florida" "Hawaii" "Maryland" "Texas" ...
##    ..$ hits    : int [1:51] 100 98 92 90 89 89 88 87 86 86 ...
##    ..$ keyword : chr [1:51] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##    ..$ geo     : chr [1:51] "US" "US" "US" "US" ...
##    ..$ gprop   : chr [1:51] "web" "web" "web" "web" ...
##  $ interest_by_dma    :'data.frame': 210 obs. of  5 variables:
##    ..$ location: chr [1:210] "Miami-Ft. Lauderdale FL" "North Platte NE" "Tampa-St. Petersburg (Saras
```

```
##   ..$ hits    : int [1:210] 100 99 98 98 97 97 96 96 96 95 ...
##   ..$ keyword : chr [1:210] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:210] "US" "US" "US" "US" ...
##   ..$ gprop   : chr [1:210] "web" "web" "web" "web" ...
##  $ interest_by_city   :'data.frame': 200 obs. of  5 variables:
##   ..$ location: chr [1:200] "Richmond West" "Pullman" "Saint Michael" "Maitland" ...
##   ..$ hits    : int [1:200] NA NA NA NA NA NA NA NA NA NA ...
##   ..$ keyword : chr [1:200] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ geo     : chr [1:200] "US" "US" "US" "US" ...
##   ..$ gprop   : chr [1:200] "web" "web" "web" "web" ...
##  $ related_topics    :'data.frame': 24 obs. of  6 variables:
##   ..$ subject       : chr [1:24] "100" "20" "7" "7" ...
##   ..$ related_topics: chr [1:24] "top" "top" "top" "top" ...
##   ..$ value         : chr [1:24] "Thanksgiving" "History" "Happiness" "Macy's Thanksgiving Day Parade
##   ..$ geo           : chr [1:24] "US" "US" "US" "US" ...
##   ..$ keyword       : chr [1:24] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ category      : int [1:24] 0 0 0 0 0 0 0 0 0 0 ...
##   ..- attr(*, "reshapeLong")=List of 4
##   .. ..$ varying:List of 1
##   .. .. ..- attr(*, "v.names")= chr "value"
##   .. .. ..- attr(*, "times")= chr "top"
##   .. ..$ v.names: chr "value"
##   .. ..$ idvar  : chr "id"
##   .. ..$ timevar: chr "related_topics"
##  $ related_queries   :'data.frame': 50 obs. of  6 variables:
##   ..$ subject       : chr [1:50] "100" "99" "83" "71" ...
##   ..$ related_topics: chr [1:50] "top" "top" "top" "top" ...
##   ..$ value         : chr [1:50] "thanksgiving history" "history of thanksgiving" "open on thanksgiv
##   ..$ geo           : chr [1:50] "US" "US" "US" "US" ...
##   ..$ keyword       : chr [1:50] "thanksgiving" "thanksgiving" "thanksgiving" "thanksgiving" ...
##   ..$ category      : int [1:50] 0 0 0 0 0 0 0 0 0 0 ...
##   ..- attr(*, "reshapeLong")=List of 4
##   .. ..$ varying:List of 1
##   .. .. ..- attr(*, "v.names")= chr "value"
##   .. .. ..- attr(*, "times")= chr "top"
##   .. ..$ v.names: chr "value"
##   .. ..$ idvar  : chr "id"
##   .. ..$ timevar: chr "related_topics"
##  - attr(*, "class")= chr [1:2] "gtrends" "list"
```

14.1.4.3 Gather Interest by State

- Select the data frame

  - for interest by state and
    * then convert the names of the states to FIPS codes
    * FIPS codes are
      · 2 characters for state,
      · 5 characters for county

The `usmap` package

- requires the data in FIPS format.

To gather the FIPS data,

- use the `usmap::fips()` function

- to convert state name to code
-

```
thanksgivingStates <- thanksgiving$interest_by_region

thanksgivingStates$fips <- fips(thanksgivingStates$location)
glimpse(thanksgivingStates$fips)
```

```
##  chr [1:51] "49" "06" "42" "34" "25" "35" "48" "17" "11" "44" "39" "33" ...
```

```
thanksgivingStates_2021 <- thanksgiving_2021$interest_by_region

thanksgivingStates_2021$fips <- fips(thanksgivingStates_2021$location)

thanksgivingStates_2020 <- thanksgiving_2020$interest_by_region

thanksgivingStates_2020$fips <- fips(thanksgivingStates_2020$location)
```

14.1.4.4 Plot interest with the US Map by state

- Create a US heatmap
  - with google search popularity
    * for the keyword "thanksgiving"

```
plot_usmap(
  data = thanksgivingStates,
  values = "hits",
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
  theme(panel.background = element_rect(colour = "black")) +
  labs(title = "Popularity of Thanksgiving this Year, Google Search by State", caption = "Source: @littl
```

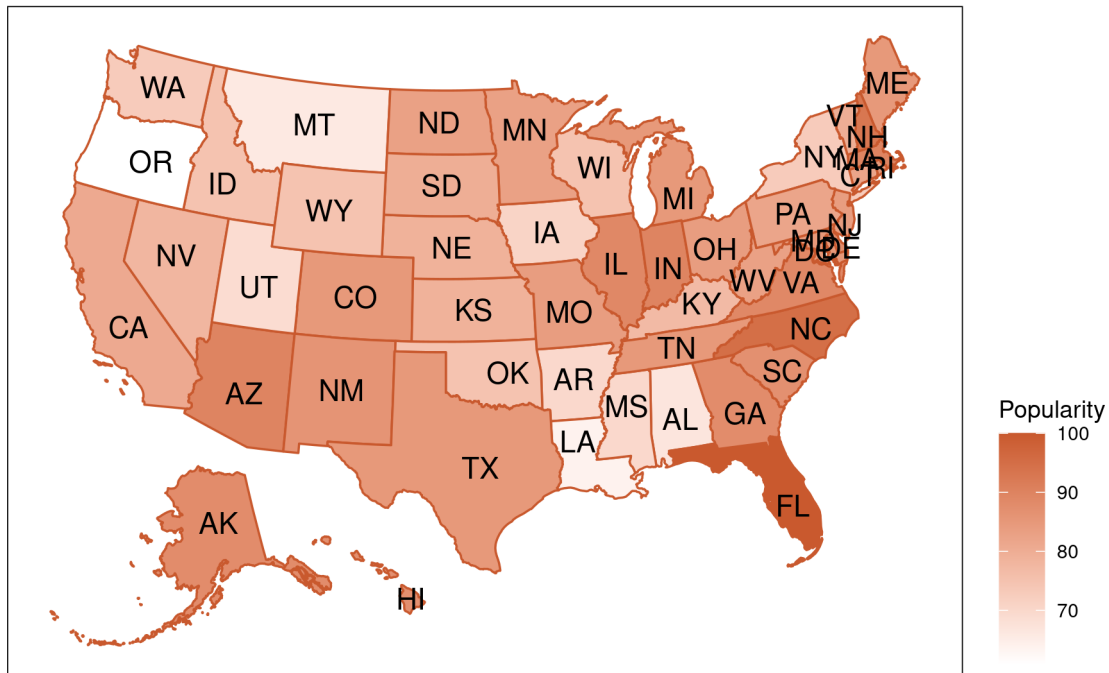Popularity of Thanksgiving this Year, Google Search by State
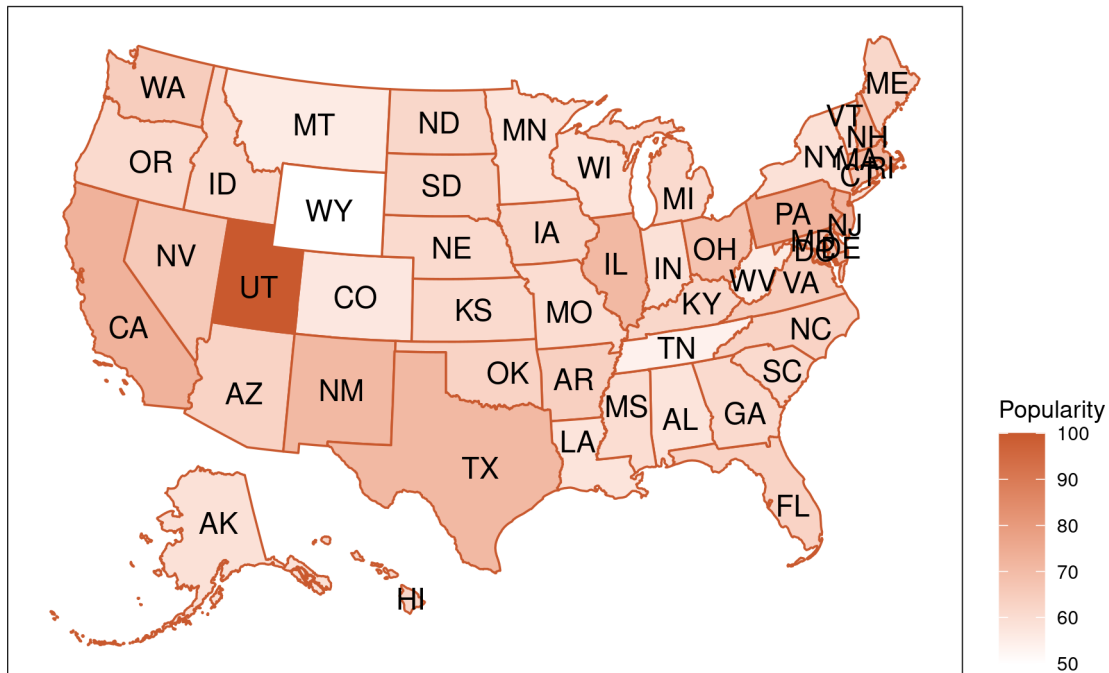
And for last year, 2020?

```
plot_usmap(
  data = thanksgivingStates_2021,
  values = "hits",
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
  theme(panel.background = element_rect(colour = "black")) +
  labs(title = "Popularity of Thanksgiving 2021, Google Search by State", caption = "Source: @littlemiss
```

Popularity of Thanksgiving 2021, Google Search by State

And for two years ago, 2020?

```
plot_usmap(
  data = thanksgivingStates_2020,
  values = "hits",
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
  theme(panel.background = element_rect(colour = "black")) +
  labs(title = "Popularity of Thanksgiving in 2020, Google Search by State", caption = "Source: @littlem
```

Popularity of Thanksgiving in 2020, Google Search by State
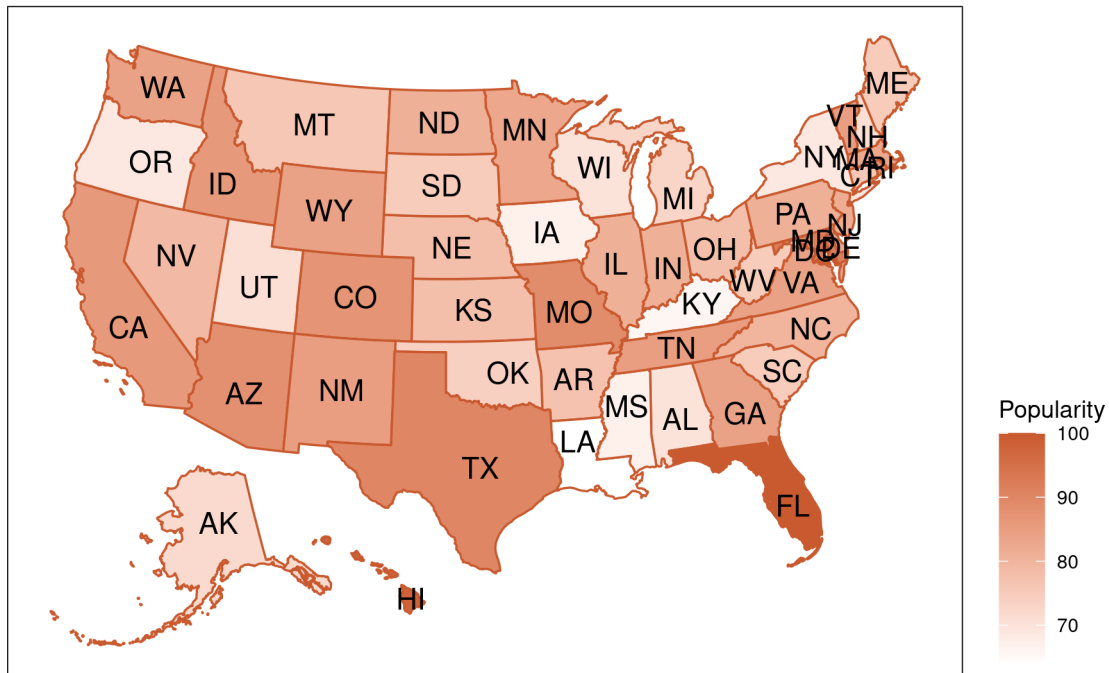
We can clearly see from the graphs above,

- differences between two years ago, last year and now.
  - about how excited people were for thanksgiving
- Lets look at the east coast
  - Now and Last Year at this time

14.1.4.5 Plot East Coast interest with the US Map by state

- Drill in on the seemingly most popular regions

 - using the "include" parameter
    * in the plot_usmap() function.

Regional divisions can be found in the docs here

For this year

```
plot_usmap(
  data = thanksgivingStates,
  values = "hits",
  include =  c(.south_atlantic, .east_north_central),
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
```

```
theme(panel.background = element_rect(colour = "black")) +
labs(title = "US East Coast Popularity of Thanksgiving 2021 Google Search", caption = "Source: @little
```

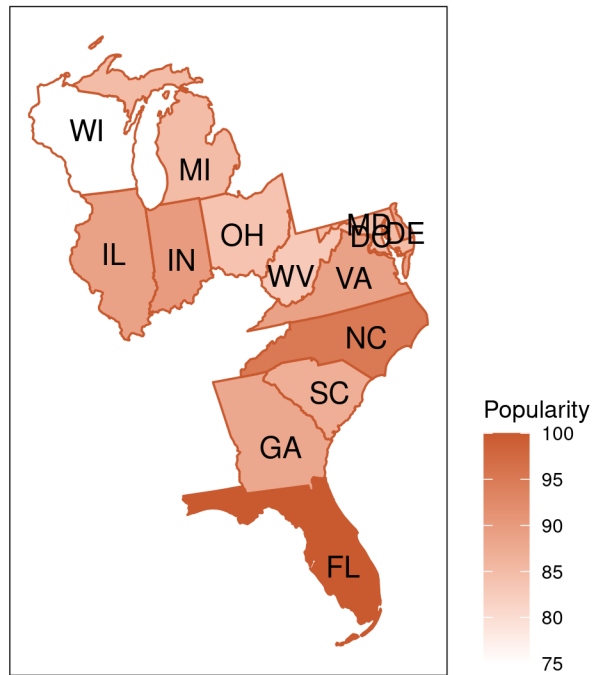## US East Coast Popularity of Thanksgiving 2021 Google Search
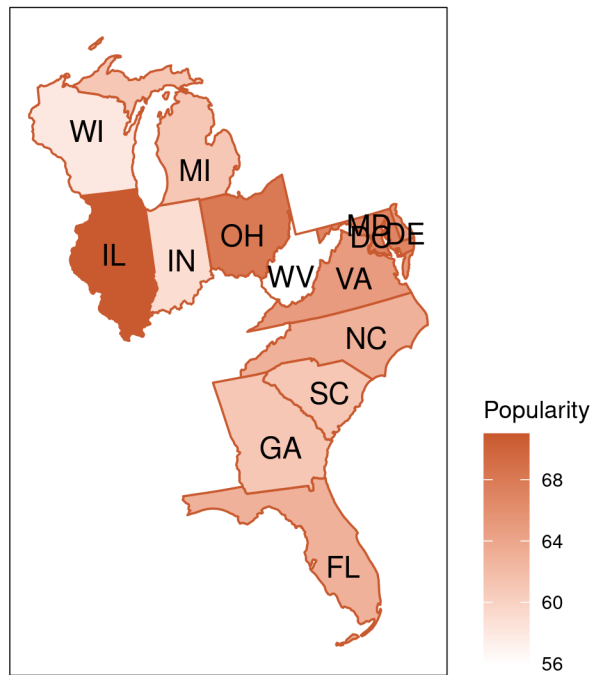
Source: @littlemissdata

And last year, 2021.

```
plot_usmap(
  data = thanksgivingStates_2021,
  values = "hits",
  include =  c(.south_atlantic, .east_north_central),
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
  theme(panel.background = element_rect(colour = "black")) +
  labs(title = "US East Coast Popularity of Thanksgiving 2021 Google Search", caption = "Source: @little
```

## US East Coast Popularity of Thanksgiving 2021 Google Search

Source: @littlemissdata

And two years ago, 2020.

```
plot_usmap(
  data = thanksgivingStates_2020,
  values = "hits",
  include =  c(.south_atlantic, .east_north_central),
  color = orange,
  labels = TRUE
) +
  scale_fill_continuous(
    low = "white",
    high = orange,
    name = "Popularity",
    label = scales::comma
  ) +
  theme(legend.position = "right") +
  theme(panel.background = element_rect(colour = "black")) +
  labs(title = "US East Coast Popularity of Thanksgiving 2020 Google Search", caption = "Source: @littl
```

US East Coast Popularity of Thanksgiving 2020 Google Search
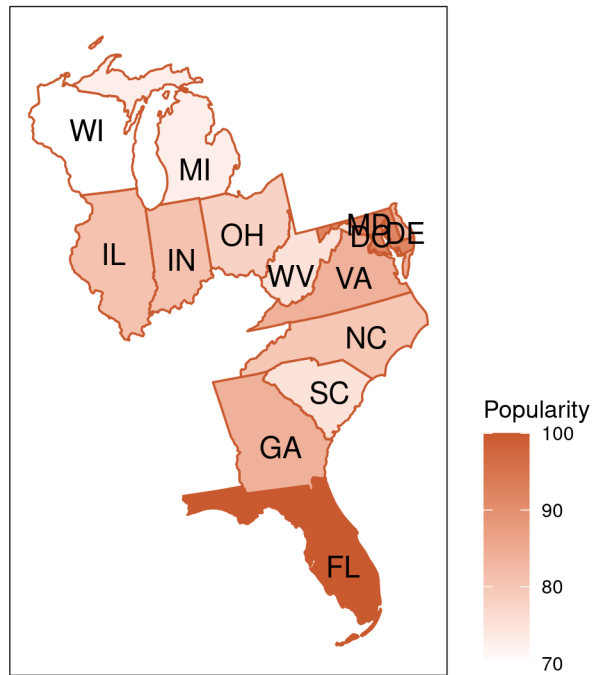
Source: @littlemissdata

14.1.5   Links

Laura Ellis, "Easy US Maps in R - Thanksgiving Edition", 11/27/2019 - https://www.littlemissdata.com/blog/usmap