

Creating Informative Maps

This was a mini-project done for SDS 264 - Data Science 2 where we are creating informative maps, 2 choropleth and 2 interactive. To view the qmd, click [here](#).

Load in Libraries & Datasets

```
# Load in libraries
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2     3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr       1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(maps)
```

Attaching package: 'maps'

The following object is masked from 'package:purrr':

map

```
library(viridis)
```

Loading required package: viridisLite

Attaching package: 'viridis'

The following object is masked from 'package:maps':

unemp

```
library(statebins)
library(leaflet)
library(sf)
```

Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE

```
library(htmltools)
library(glue)
library(readr)
```

```
# Load in datasets
```

```
decennial_census10 <- read_csv("~/SDS264/Class Files/00_Data/DECENNIALCD1162010.P1-Data.csv")
```

New names:

Rows: 53 Columns: 4

-- Column specification

----- Delimiter: "," chr

(3): GEO_ID, NAME, P001001 lgl (1): ...4

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.

* `` -> `...4`

```
decennial_census20 <- read_csv("~/SDS264/Class Files/00_Data/DECENNIALCD1182020.P1-Data.csv")
```

New names:

Rows: 53 Columns: 4

-- Column specification

----- Delimiter: "," chr

(3): GEO_ID, NAME, P1_001N lgl (1): ...4

```
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.
* `` -> `...4`
```

```
marijuana_laws2024 <- read_csv("~/SDS264/Class Files/00_Data/marijuana-laws-by-state-2024.csv")
```

```
Rows: 51 Columns: 4
-- Column specification -----
Delimiter: ","
chr (4): state, LegalStatus, Medicinal, Decriminalized
```

```
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.
```

```
# Load in states sf
states <- read_sf("https://rstudio.github.io/leaflet/json/us-states.geojson") ②
states
```

Simple feature collection with 52 features and 3 fields

Geometry type: MULTIPOLYGON

Dimension: XY

Bounding box: xmin: -188.9049 ymin: 17.92956 xmax: -65.6268 ymax: 71.35163

Geodetic CRS: WGS 84

A tibble: 52 x 4

	id	name	density	geometry
	<chr>	<chr>	<dbl>	<MULTIPOLYGON [°]>
1	01	Alabama	94.6	(((−87.3593 35.00118, −85.60667 34.98475~
2	02	Alaska	1.26	(((−131.602 55.11798, −131.5692 55.28229~
3	04	Arizona	57.0	(((−109.0425 37.00026, −109.048 31.33163~
4	05	Arkansas	56.4	(((−94.47384 36.50186, −90.15254 36.4963~
5	06	California	242.	(((−123.2333 42.00619, −122.3789 42.0116~
6	08	Colorado	49.3	(((−107.9197 41.00391, −105.729 40.99843~
7	09	Connecticut	739.	(((−73.05353 42.03905, −71.79931 42.0226~
8	10	Delaware	464.	(((−75.41409 39.80446, −75.5072 39.68396~
9	11	District of Columbia	10065	(((−77.03526 38.99387, −76.90929 38.8952~
10	12	Florida	353.	(((−85.49714 30.99754, −85.00421 31.0030~

i 42 more rows

Joining Datasets & Calculating Percent Change

```
# Data wrangling
census_sf <- decennial_census10 |>
  left_join(decennial_census20, join_by(NAME)) |>
  rename(name = NAME,
         population20 = P1_001N,
         population10 = P001001) |>
  select(name,
         population10,
         population20) |>
  filter(!(name %in% c("Alaska",
                     "Hawaii",
                     "District of Columbia",
                     "Puerto Rico",
                     "Label for GEO_ID")))
) |>
  mutate(population10 = as.double(population10),
         population20 = as.double(population20),
         perc_change = (population20 - population10)/abs(population10) * 100,
         perc_change = round(perc_change, 3)) |>
  left_join(states, join_by(name)) |>
  st_as_sf()

census_sf
```

Simple feature collection with 48 features and 6 fields

Geometry type: MULTIPOLYGON

Dimension: XY

Bounding box: xmin: -124.7066 ymin: 25.12078 xmax: -66.9796 ymax: 49.38362

Geodetic CRS: WGS 84

A tibble: 48 x 7

	name	population10	population20	perc_change	id	density
	<chr>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
1	Alabama	4779736	5024279	5.12	01	94.6
2	Arizona	6392017	7151502	11.9	04	57.0
3	Arkansas	2915918	3011524	3.28	05	56.4
4	California	37253956	39538223	6.13	06	242.
5	Colorado	5029196	5773714	14.8	08	49.3
6	Connecticut	3574097	3605944	0.891	09	739.
7	Delaware	897934	989948	10.2	10	464.

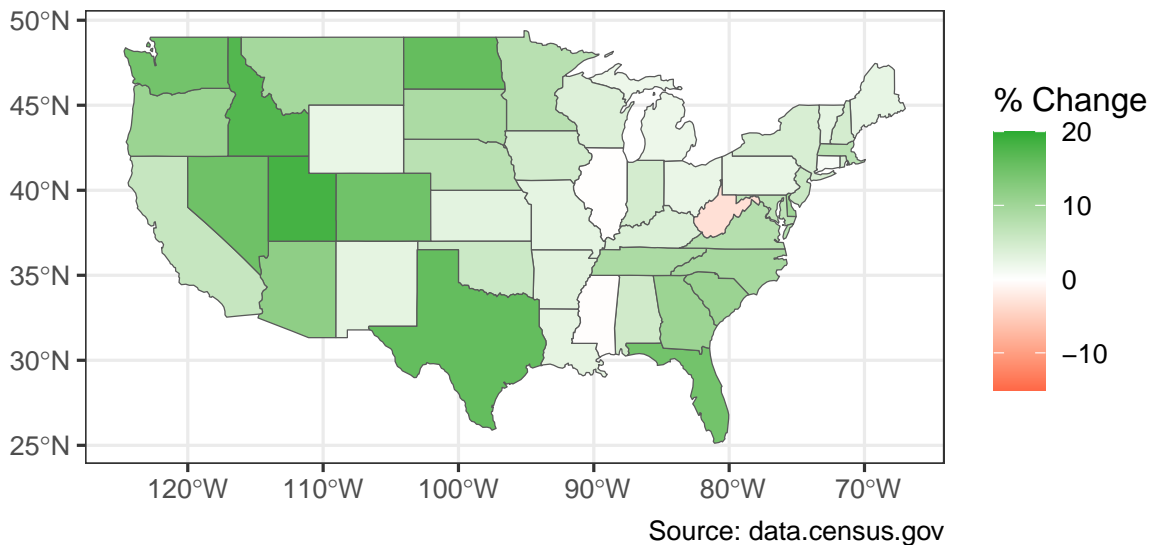
8	Florida	18801310	21538187	14.6	12	353.
9	Georgia	9687653	10711908	10.6	13	170.
10	Idaho	1567582	1839106	17.3	16	19.2

i 38 more rows
i 1 more variable: geometry <MULTIPOLYGON [°]>

Percentage Change in Population of the US from 2010 to 2020

```
# First static plot
ggplot(census_sf) +
  geom_sf(aes(fill = perc_change)) +
  scale_fill_gradient2(
    high = "#2aab32",
    mid = "white",
    low = "red",
    limits = c(-15, 20)) +
  labs(title = "Percentage Change in Population of the US from 2010 to 2020",
       fill = "% Change",
       caption = " Source: data.census.gov") +
  theme_bw()
```

Percentage Change in Population of the US from 2010 to 2020



This is a map of the United States showing the percent change in population of the United States from 2010 to 2020 by state. On the y-axis is longitude and on the x-axis is latitude. Percent change varies from a decrease of -3% of the population to an increase of 18%, with red representing a decrease in population (a negative percentage) and green representing an increase in population (a positive percentage). We can see from this map that Utah, Idaho, and Texas had the largest percent increase in population. The states in the Midwest, such as Illinois, Iowa, Minnesota, South Dakota, Wisconsin, etc. (the 90W and 45N area), the states also had an increase in population however it is a smaller percentage. West Virginia is the only visible state that is red and this represents a decrease in population by 3.19%. For the population increase in states, this could be caused by people moving states, and more people in the United States overall.

Same Map, but Interactive!

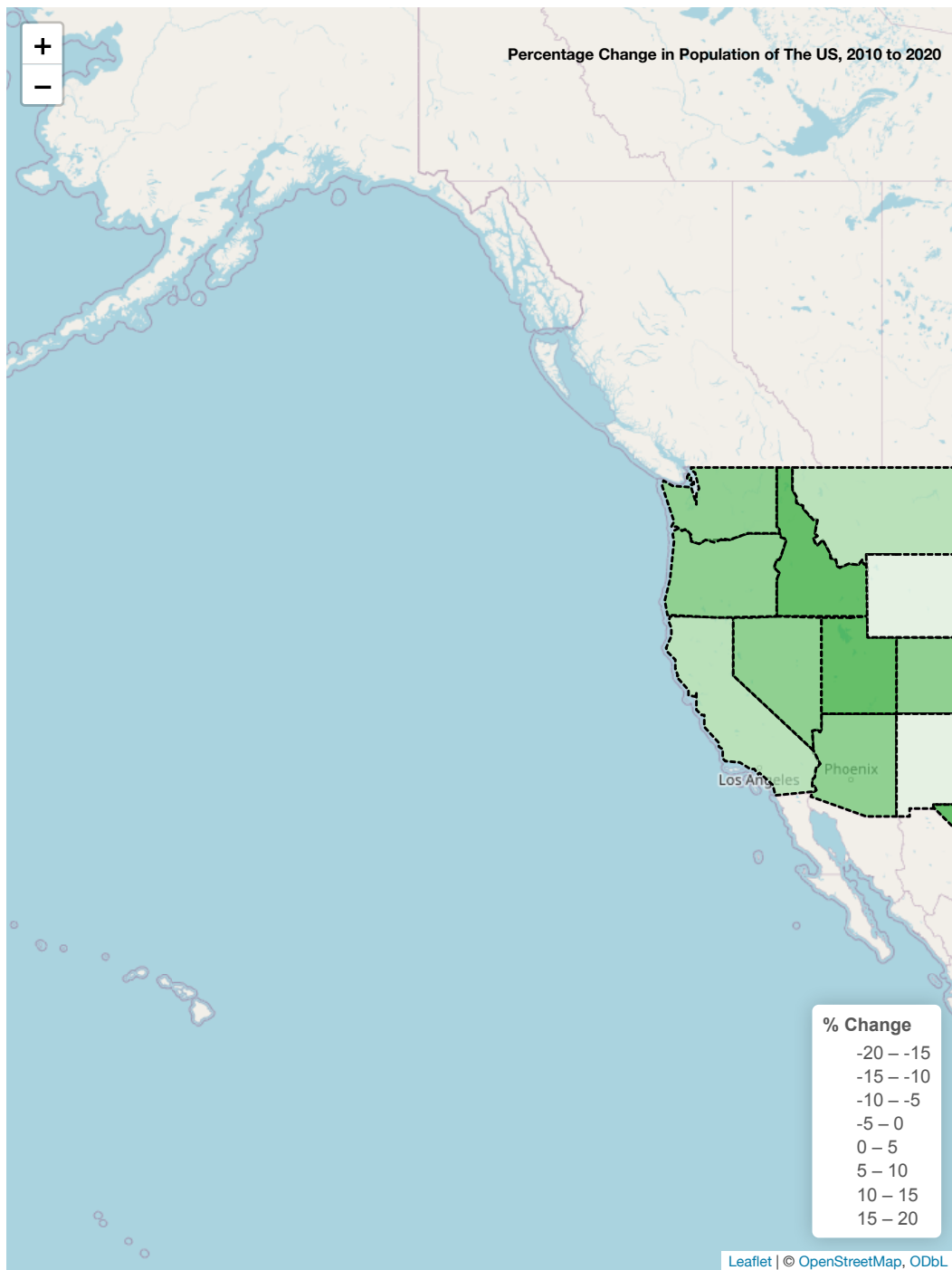
```
# First interactive plot
bins <- c(-20, -15, -10, -5, 0, 5, 10, 15, 20)
my_scale <- colorRamp(c("red", "white", "#2aab32"), interpolate="linear")
pal <- colorBin(palette = my_scale, domain = census_sf$perc_change, bins = bins)
```

```

census_sf <- census_sf |>
  mutate(labels = str_c(name, ": ", perc_change, "% Change in Population"))
  labels <- lapply(census_sf$labels, HTML)

leaflet(census_sf) |>
  setView(-96, 37.8, 4) |>
  addTiles() |>
  addPolygons(
    weight = 2,
    opacity = 1,
    color = "black",
    fillColor = ~ pal(census_sf$perc_change),
    dashArray = "3",
    fillOpacity = 0.7,
    highlightOptions = highlightOptions(
      weight = 5,
      color = "#666",
      dashArray = "",
      fillOpacity = 0.7,
      bringToFront = TRUE),
    label = labels,
    labelOptions = labelOptions(
      style = list("font-weight" = "normal", padding = "3px 8px"),
      textSize = "15px",
      direction = "auto")) |>
  addLegend(pal = pal,
    values = ~census_sf$perc_change,
    opacity = 0.7,
    title = "% Change",
    position = "bottomright") |>
  addControl("<h4>Percentage Change in Population of The US, 2010 to 2020</h4>",
    position = "topright",
    className = "map-title")

```



Creating New “legality” Column

```
# Data wrangling pt. 2
marijuana_laws2024_sf <- marijuana_laws2024 |>
  mutate(legality = ifelse(LegalStatus == "Mixed" &
    Medicinal == "Yes" &
    Decriminalized == "No", "Medicinal, not decriminalized", LegalStatus),
    legality = ifelse(LegalStatus == "Mixed" &
    Medicinal == "Yes" &
    Decriminalized == "Yes", "Medicinal, decriminalized", legality),
    legality = ifelse(LegalStatus == "Mixed" &
    Medicinal == "CBD Oil Only" &
    Decriminalized == "No", "CBD Oil Only, not decriminalized", legality),
    legality = ifelse(LegalStatus == "Mixed" &
    Medicinal == "CBD Oil Only" &
    Decriminalized == "Yes", "CBD Oil Only, decriminalized", legality),
    legality = ifelse(LegalStatus == "Legalized" &
    Decriminalized == "Yes", "Legalized, decriminalized", legality),
    legality = ifelse(LegalStatus == "Legalized" &
    Decriminalized == "No", "Legalized, not decriminalized", legality),
    legality = ifelse(LegalStatus == "Fully Illegal" &
    Decriminalized == "No", "Illegal, not decriminalized", legality),
    legality = ifelse(LegalStatus == "Fully Illegal" &
    Decriminalized == "Yes", "Illegal, decriminalized", legality),
    legality = ifelse(LegalStatus == "Illegal, but decriminalized" &
    Decriminalized == "Yes", "Illegal, decriminalized", legality),
    LegalStatus = ifelse(LegalStatus == "Fully Illegal", "Illegal", LegalStatus),
    LegalStatus = ifelse(LegalStatus == "Illegal, but decriminalized", "Illegal", LegalStatus)
  ) |>
  rename(name = state) |>
  filter(!(name %in% c("Alaska", "Hawaii", "District of Columbia", "Puerto Rico"))) |>
  left_join(states, join_by(name)) |>
  st_as_sf()

marijuana_laws2024_sf
```

Simple feature collection with 48 features and 7 fields

Geometry type: MULTIPOLYGON

Dimension: XY

Bounding box: xmin: -124.7066 ymin: 25.12078 xmax: -66.9796 ymax: 49.38362

Geodetic CRS: WGS 84

```
# A tibble: 48 x 8
```

	name <chr>	LegalStatus <chr>	Medicinal <chr>	Decriminalized <chr>	legality <chr>	id <chr>	density <dbl>
1	Alabama	Mixed	Yes	No	Medicinal,~	01	94.6
2	Arizona	Legalized	Yes	Yes	Legalized,~	04	57.0
3	Arkansas	Mixed	Yes	No	Medicinal,~	05	56.4
4	California	Legalized	Yes	Yes	Legalized,~	06	242.
5	Colorado	Legalized	Yes	Yes	Legalized,~	08	49.3
6	Connecticut	Legalized	Yes	Yes	Legalized,~	09	739.
7	Delaware	Legalized	Yes	Yes	Legalized,~	10	464.
8	Florida	Mixed	Yes	No	Medicinal,~	12	353.
9	Georgia	Mixed	CBD Oil Only	No	CBD Oil On~	13	170.
10	Idaho	Illegal	No	No	Illegal, n~	16	19.2

```
# i 38 more rows
```

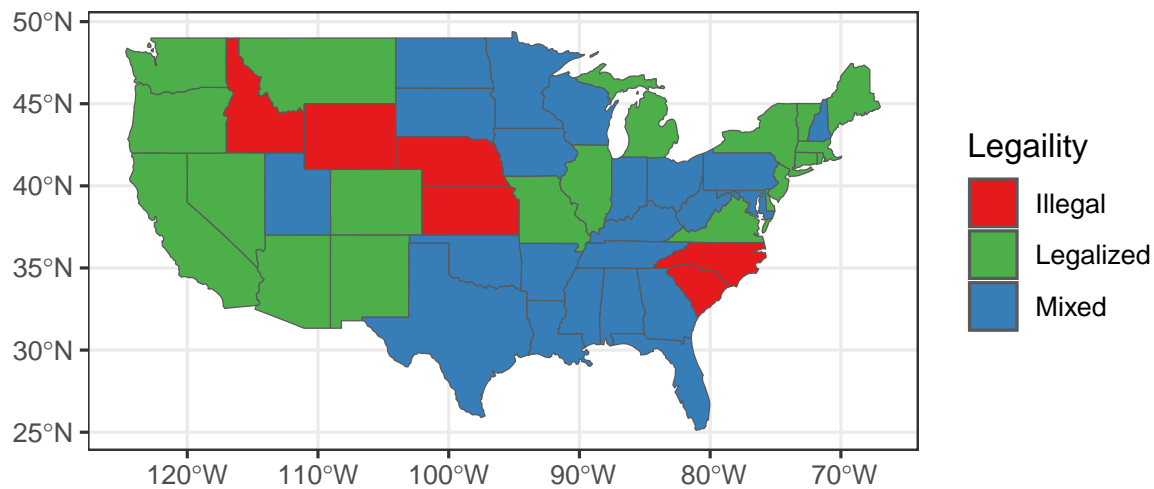
```
# i 1 more variable: geometry <MULTIPOLYGON [°]>
```

Legalization of Marijuana in the US 2022

```
# Second static plot
```

```
ggplot(marijuana_laws2024_sf) +  
  geom_sf(aes(fill = LegalStatus)) +  
  scale_fill_manual(values = c("#e41a1c", "#4daf4a", "#377eb8")) +  
  labs(title = "Legalization of Marijuana in the US 2022",  
        fill = "Legaility",  
        caption = "Source: https://worldpopulationreview.com/state-rankings/marijuana-laws-by",  
        theme_bw()
```

Legalization of Marijuana in the US 2022



Source: <https://worldpopulationreview.com/state-rankings/marijuana-laws-by-state>

This is a map of the United States showing the legalization of marijuana in 2022. On the y-axis is longitude and on the x-axis is latitude. The categories of legality are “Illegal”, “Legalized”, and “Mixed”, with the associated colors on the map being red, green, and blue, respectively. There are 6 states where marijuana is illegal, these states are Idaho, Kansas, Nebraska, North and South Carolina, and Wyoming. On the map, 2 of these states are on the East Coast and the rest are in the middle of the US. For the rest of the US, marijuana is either legalized or has mixed laws. A lot of the states in the west have legalized marijuana, this includes Oregon, Washington, California, along with Nevada and Arizona. There are also some states on the East Coast have also legalized marijuana. Whereas states in the south and in the Midwest have mixed laws. Note that this is data from 2022, and some states with mixed laws, are now fully legalized.

Same Map, but Interactive!

```
# Second interactive plot
marijuana_laws2024_sf <- marijuana_laws2024_sf |>
  mutate(labels = str_c(name, ":", " ", legality))

labels <- lapply(marijuana_laws2024_sf$labels, HTML)
```

```

factpal <- colorFactor(c("#e41a1c", "#4daf4a", "#377eb8"),
                      levels(marijuana_laws2024_sf$LegalStatus))

leaflet(marijuana_laws2024_sf) |>
  setView(-96, 37.8, 4) |>
  addTiles() |>
  addPolygons(
    weight = 2,
    opacity = 1,
    color = "black",
    fillColor = ~ factpal(marijuana_laws2024_sf$LegalStatus),
    dashArray = "3",
    fillOpacity = 0.7,
    highlightOptions = highlightOptions(
      weight = 5,
      color = "#666",
      dashArray = "",
      fillOpacity = 0.7,
      bringToFront = TRUE),
    label = labels,
    labelOptions = labelOptions(
      style = list("font-weight" = "normal",
                  padding = "3px 8px"),
      textsize = "15px",
      direction = "auto")) |>
  addLegend(pal = factpal,
            values = ~marijuana_laws2024_sf$LegalStatus,
            opacity = 0.7, title = NULL,
            position = "bottomright") |>
  addControl("<h4>Legalization of Marijuana in the US 2022</h4>",
            position = "topright",
            className = "map-title")

```

