

Homework 7 - Geoplotting US Wage Data

References:

- https://map-rfun.library.duke.edu/032_thematic_mapping_geom_sf.html
- <https://www.computerworld.com/article/3175623/mapping-in-r-just-got-a-whole-lot-easier.html>

In this homework, we plot the mean wages over the US states.

The purpose of this analysis is to visualize the average wages for the different US states.

The Data

```
library(tidyverse)
library(readxl)
library(tigris)
library(sf)
library(viridis)
```

The data for this homework come from two sources.

1. The `tigris` dataset provides the state geographies.

TODO: This code chunk doesn't work on rstudio.calvin.edu for me, regardless of the versions of tidyverse, sf, tigris being used. Two work-arounds - use a local RStudio or drop the `shift_geometry()` call entirely.

```
us_state_geographies <- tigris::states(cb=TRUE, resolution="20m") %>%
  shift_geometry()
```

Retrieving data for the year 2021

```
us_state_geographies
```

2. This spreadsheet provides the mean wages by state. You'll need to cleanse this data.

```
state_mean_wages_raw <-
  read_excel("data/BLS-wages-by-state.xlsx",
             col_types = c("text", "numeric"),
             skip = 4)
state_mean_wages_raw
```

```
state_mean_wages <- state_mean_wages_raw %>%
  rename(wages = "Annual mean wage(2)") %>%
  mutate(State = str_extract(`Area Name`, "\\w+.*(?=\\(\\))") %>%
    drop_na(wages) %>%
    select(State, wages)
state_mean_wages
```

```
## # A tibble: 53 x 2
##   State      wages
##   <chr>      <dbl>
## 1 Alabama    38080
## 2 Alaska     46050
## 3 Arizona    34560
## 4 Arkansas   40480
## 5 California 64320
## 6 Colorado   43480
## 7 Connecticut 60450
## 8 Delaware   54450
## 9 District of Columbia 57210
## 10 Florida    43410
## # ... with 43 more rows
```

Data Wrangling

You'll combine these datasets to produce geospatially-categorized wage data, and focus only on the rows for the 50 states.

```
us_state_geographies_wages <- us_state_geographies %>%
  left_join(state_mean_wages,
    by = c("NAME" = "State"))
us_state_geographies_wages
```

```
## Simple feature collection with 52 features and 10 fields
## Geometry type: MULTIPOLYGON
## Dimension: XY
## Bounding box: xmin: -3112200 ymin: -1697728 xmax: 2258154 ymax: 1558935
## Projected CRS: USA_Contiguous_Albers_Equal_Area_Conic
## First 10 features:
##   STATEFP STATENS AFFGEOID GEOID STUSPS NAME LSAD ALAND
## 1 22 01629543 0400000US22 22 LA Louisiana 00 111915258185
## 2 24 01714934 0400000US24 24 MD Maryland 00 25151992308
## 3 55 01779806 0400000US55 55 WI Wisconsin 00 140292321657
## 4 12 00294478 0400000US12 12 FL Florida 00 138961722096
## 5 13 01705317 0400000US13 13 GA Georgia 00 149486624386
## 6 47 01325873 0400000US47 47 TN Tennessee 00 106791957894
## 7 27 00662849 0400000US27 27 MN Minnesota 00 206232627084
## 8 19 01779785 0400000US19 19 IA Iowa 00 144659721012
## 9 29 01779791 0400000US29 29 MO Missouri 00 178052333716
## 10 26 01779789 0400000US26 26 MI Michigan 00 146614937132
##   AWATER wages geometry
## 1 23736382213 37270 MULTIPOLYGON (((182432.3 -5...
```

```
## 2 6979074857 47560 MULTIPOLYGON (((1722285 240...
## 3 29343646672 46220 MULTIPOLYGON (((708320.1 91...
## 4 45972570361 43410 MULTIPOLYGON (((1451445 -13...
## 5 4418360134 52750 MULTIPOLYGON (((939223.1 -2...
## 6 2322601737 35540 MULTIPOLYGON (((515382.5 -2...
## 7 18949394733 49310 MULTIPOLYGON (((-91052.17 1...
## 8 1085964740 45380 MULTIPOLYGON (((-50588.83 5...
## 9 2487445715 35510 MULTIPOLYGON (((19670.04 34...
## 10 103871870536 49420 MULTIPOLYGON (((882371.5 99...
```

```
contiguous_states <- us_state_geographies_wages %>%
  filter(GEOID < 60)
contiguous_states
```

```
## Simple feature collection with 51 features and 10 fields
## Geometry type: MULTIPOLYGON
## Dimension: XY
## Bounding box: xmin: -3112200 ymin: -1697728 xmax: 2258154 ymax: 1558935
## Projected CRS: USA_Contiguous_Albers_Equal_Area_Conic
## First 10 features:
## STATEFP STATENS AFFGEOID GEOID STUSPS NAME LSAD ALAND
## 1 22 01629543 0400000US22 22 LA Louisiana 00 111915258185
## 2 24 01714934 0400000US24 24 MD Maryland 00 25151992308
## 3 55 01779806 0400000US55 55 WI Wisconsin 00 140292321657
## 4 12 00294478 0400000US12 12 FL Florida 00 138961722096
## 5 13 01705317 0400000US13 13 GA Georgia 00 149486624386
## 6 47 01325873 0400000US47 47 TN Tennessee 00 106791957894
## 7 27 00662849 0400000US27 27 MN Minnesota 00 206232627084
## 8 19 01779785 0400000US19 19 IA Iowa 00 144659721012
## 9 29 01779791 0400000US29 29 MO Missouri 00 178052333716
## 10 26 01779789 0400000US26 26 MI Michigan 00 146614937132
## AWATER wages geometry
## 1 23736382213 37270 MULTIPOLYGON (((182432.3 -5...
## 2 6979074857 47560 MULTIPOLYGON (((1722285 240...
## 3 29343646672 46220 MULTIPOLYGON (((708320.1 91...
## 4 45972570361 43410 MULTIPOLYGON (((1451445 -13...
## 5 4418360134 52750 MULTIPOLYGON (((939223.1 -2...
## 6 2322601737 35540 MULTIPOLYGON (((515382.5 -2...
## 7 18949394733 49310 MULTIPOLYGON (((-91052.17 1...
## 8 1085964740 45380 MULTIPOLYGON (((-50588.83 5...
## 9 2487445715 35510 MULTIPOLYGON (((19670.04 34...
## 10 103871870536 49420 MULTIPOLYGON (((882371.5 99...
```

Data Visualization

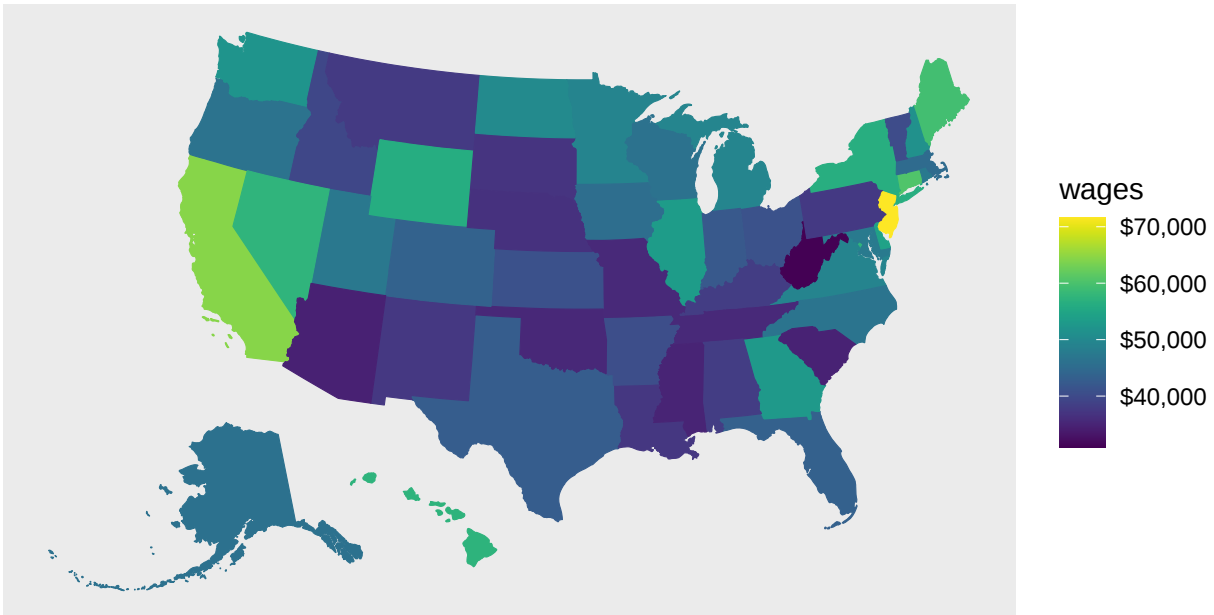
Produce a choropleth visualization that shows the average annual wages for all 50 states.

Do you notice any interesting patterns in this data?

```
contiguous_states %>%
  ggplot(aes(fill = wages, color = wages)) +
  geom_sf() +
```

```
coord_sf(crs = 5070, datum = NA) +
scale_fill_viridis(label = scales::dollar) +
scale_color_viridis(label = scales::dollar) +
labs(title = "Annual Mean Wages by State",
      caption = "Data Source: BLS.gov - Occupational Employment Statistics ; 2016")
```

Annual Mean Wages by State



Data Source: BLS.gov – Occupational Employment Statistics ; 2016

High median salaries generally appear in the southwest and northeast. West Virginia is shockingly low, and other states in the mid-south are low as well.