Maps vignette - Geo-location per country

ggplot geom_sf() and rnaturalearth

Oscar A. Trevizo

2023-05-10

Contents

1 Documentation

From > ?geom_sf() documentation:

> "This set of geom, stat, and coord are used to visualise simple feature (sf) objects. For simple plots, you will only need geom_sf() as it uses stat_sf() and adds coord_sf() for you. geom_sf() is an unusual geom because it will draw different geometric objects depending on what simple features are present in the data: you can get points, lines, or polygons. For text and labels, you can use geom_sf_text() and geom_sf_label()."

We need to get world polygons from *rnaturalearth*.

2 Load libraries

3 Load the data:

```
# Based on migration project from Harvard Statistics course (May 2023)
wpp_wb <- read.csv('../data/wpp_wb_2023.Q1.csv', check.names=FALSE)
wpp_wb$mig_pattern <- as.factor(wpp_wb$mig_pattern)
wpp_wb$migration_swings <- as.factor(wpp_wb$migration_swings)
wpp_wb$emigrates <- as.factor(wpp_wb$emigrates)
```

4 Get countries average coordinates

We need to map average long / lat coordinates per country. The average coordinates will be needed to plug in the bubbles per country.

```
# Reference https://developers.google.com/public-data/docs/canonical/countries_csv
# Reference (has dups) https://gist.github.com/tadast/8827699
# Reference (w/o dups) https://gist.github.com/cpl/3dc2d19137588d9ae202d67233715478

# Here I downloaded the last reference w/o dups
w <- read.csv('../data/countries_codes_and_coordinates.csv')

# Rename some variable ahead of merging the dataset with another dataset.
w <- w %>% rename(country = Country,
                          ISO2 = Alpha.2.code,
                          ISO3 = Alpha.3.code,
                          group = Numeric.code,
                          lat = Latitude..average.,
                          long = Longitude..average.)

# Simplify
w <- w %>% dplyr::select(ISO3, group, lat, long)

# Strip leading spaces
# https://www.geeksforgeeks.org/remove-all-whitespace-in-each-dataframe-column-in-r/
w <- as.data.frame(apply(w, 2, function(x) gsub("\\s+", "", x)))

# Typsets
w[, 3:4] <- sapply(w[, 3:4], as.numeric)

head(w)
```

```
##   ISO3 group      lat   long
## 1  AFG     4  33.0000   65.0
## 2  ALB     8  41.0000   20.0
## 3  DZA    12  28.0000    3.0
## 4  ASM    16 -14.3333 -170.0
## 5  AND    20  42.5000    1.6
## 6  AGO    24 -12.5000   18.5
```

5 Get country polygons from rnaturalearth

The polygons are necessary to draw each country in a map.

```
world <- ne_countries(scale = "medium", returnclass = "sf")
```

6 EDA a world dataframe

```
# Step 1, get average long lat per country
wpp_geo <- wpp_wb %>% group_by(ISO3, country, mig_pattern, net_mig_rate_med) %>%
  summarize(net_mig_mu = mean(net_migrants, na.rm = TRUE),
            net_mig_me = mean(net_migrants, na.rm = TRUE)) %>%
  left_join(w, by = join_by(ISO3))
```

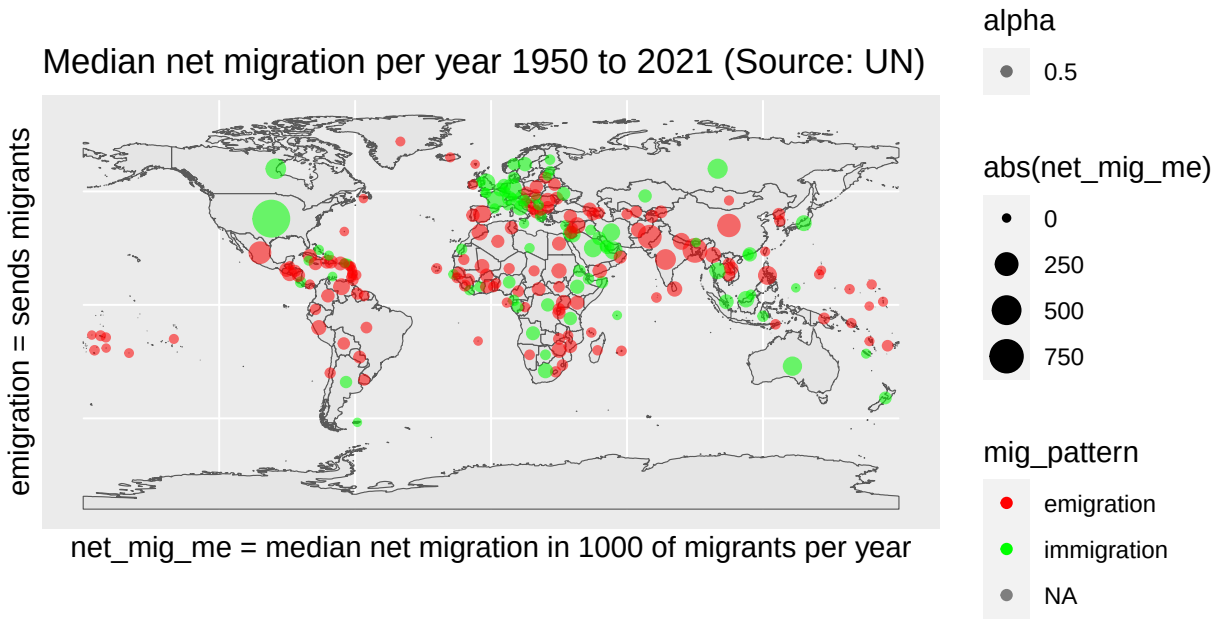
```
## `summarise()` has grouped output by 'ISO3', 'country', 'mig_pattern'. You can
## override using the `.groups` argument.
```

```
world <- left_join(world, wpp_geo, join_by('iso_a3' == 'ISO3'))
```

7 Bubbles: Geoplot med net migration with migration pattern

```
world %>% ggplot() +
  geom_sf() +
  geom_point(aes( x = `long`,
                  y = `lat`,
                  size = abs(net_mig_me),
                  alpha = 0.5,
                  color = mig_pattern)) +
  scale_color_discrete(type = c('red', 'green')) +
  labs(title = "Median net migration per year 1950 to 2021 (Source: UN)") +
  xlab('net_mig_me = median net migration in 1000 of migrants per year') +
  ylab('emigration = sends migrants')
```

```
## Warning: Removed 19 rows containing missing values (`geom_point()`).
```

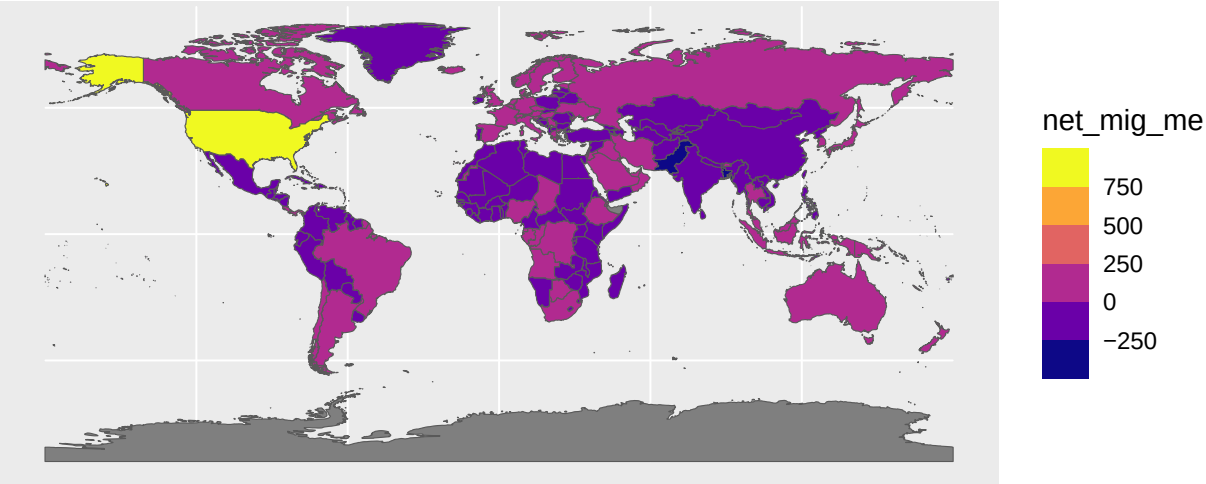


8 Filled: Geoplot med net migration

```
# https://ggplot2.tidyverse.org/reference/scale_viridis.html
world %>% ggplot() + geom_sf(mapping = aes(fill = net_mig_me))  +
  scale_fill_viridis_b(option = 'plasma') +
  labs(title = "Med Net Migrants (thousands) 1950 to 2021",
  subtitle = "Source: United Nations World Population Prospects") +
  xlab('net_mig_me = median net migration in 1000 of migrants per year')
```

Med Net Migrants (thousands) 1950 to 2021

Source: United Nations World Population Prospects



net_mig_me = median net migration in 1000 of migrants per year