Project 2

Jeremy Rindal

06 April, 2023

```
knitr::opts_chunk$set(echo = TRUE,
                      fig.align = 'center')
```

1. Imports and cleaning the data

```
library(tidyverse)    # For ggplot, dplyr, and friends
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## 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(countrycode)  # For dealing with country names, abbreviations, and codes
library(lubridate)    # For dealing with dates
```

```
## Loading required package: timechange
##
## Attaching package: 'lubridate'
##
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(WDI)

# Reading in csv file from local path
refugees_raw <- read_csv("/Users/jeremy/Desktop/MSU Classes/SSC 442/Project 2/refugee_status.csv", na =
```

```
## Rows: 70 Columns: 11
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (1): Continent/Country of Nationality
## num (10): 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015
##
## 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.
```

```
# Cleaning the data
non_countries <- c("Africa", "Asia", "Europe", "North America", "Oceania",
                   "South America", "Unknown", "Other", "Total")
```

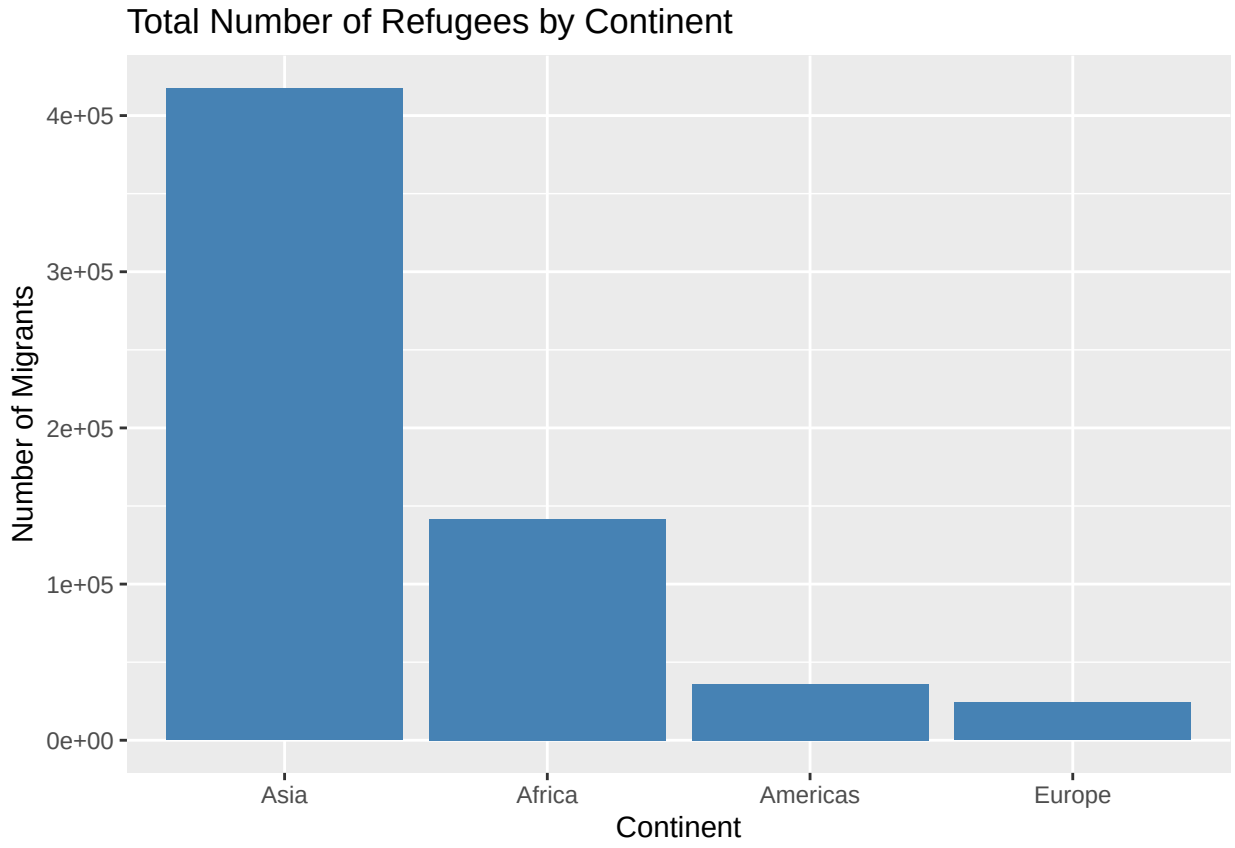
```r
refugees_clean <- refugees_raw %>%
  # Make this column name easier to work with
  rename(origin_country = `Continent/Country of Nationality`) %>%
  # Get rid of non-countries
  filter(!(origin_country %in% non_countries)) %>%
  # Convert country names to ISO3 codes
  mutate(iso3 = countrycode(origin_country, "country.name", "iso3c",
                            custom_match = c("Korea, North" = "PRK"))) %>%
  # Convert ISO3 codes to country names, regions, and continents
  mutate(origin_country = countrycode(iso3, "iso3c", "country.name"),
         origin_region = countrycode(iso3, "iso3c", "region"),
         origin_continent = countrycode(iso3, "iso3c", "continent")) %>%
  # Make this data tidy
  gather(year, number, -origin_country, -iso3, -origin_region, -origin_continent) %>%
  # Make sure the year column is numeric + make an actual date column for years
  mutate(year = as.numeric(year),
         year_date = ymd(paste0(year, "-01-01")))
```

2. Summarizing the data

```
# Summing continent totals
summary_by_continent <- aggregate(number ~ origin_continent, data = refugees_clean, FUN = sum)
# Viewing results
summary_by_continent
```

```
##   origin_continent number
## 1           Africa 141682
## 2         Americas  36187
## 3             Asia 417479
## 4           Europe  24119
```

```
ggplot(summary_by_continent, aes(x = reorder(origin_continent, -number), y = number)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  labs(title = "Total Number of Refugees by Continent", x = "Continent", y = "Number of Migrants")
```

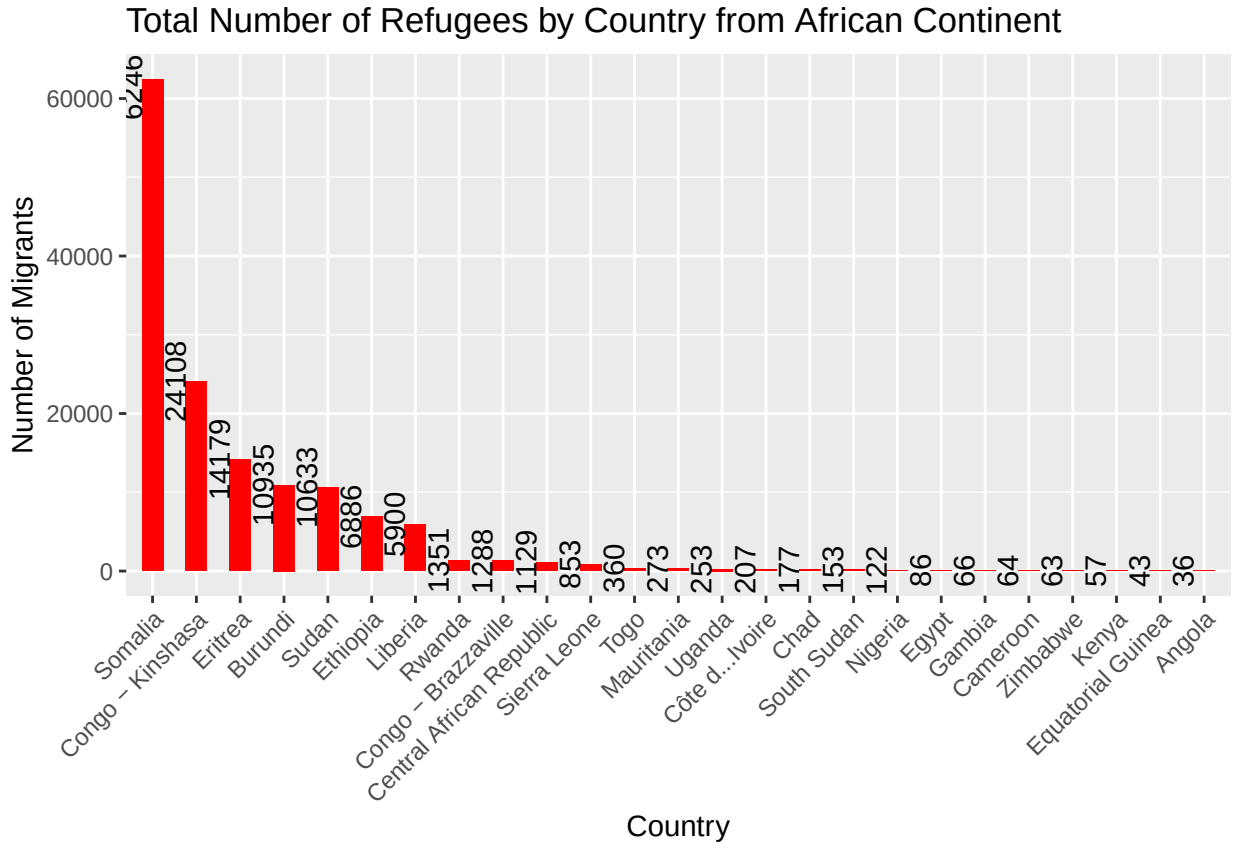


```r
africa_refugees <- refugees_clean[refugees_clean$origin_continent == "Africa", ]
africa_summary <- aggregate(number ~ origin_country, data = africa_refugees, FUN = sum)
africa_summary <- africa_summary[africa_summary$number > 1, ]

ggplot(africa_summary, aes(x = reorder(origin_country, -number), y = number)) +
  geom_bar(stat = "identity", fill = "red", width = 0.5) +
  geom_text(aes(label = number), vjust = -0.5, angle = 90) + # rotate the bar labels
  labs(title = "Total Number of Refugees by Country from African Continent", x = "Country", y = "Number
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

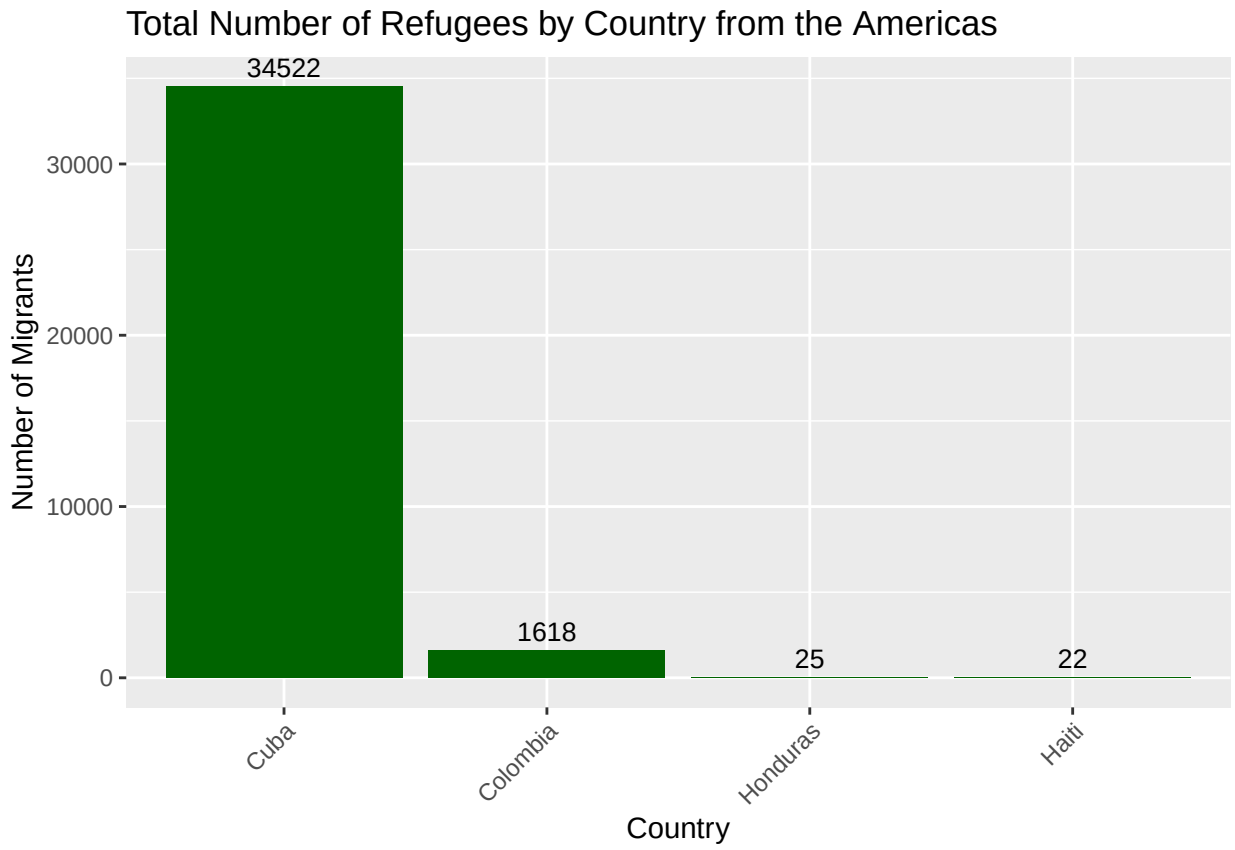
```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d'Ivoire' in 'mbcsToSbcs': dot substituted for <80>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>
```

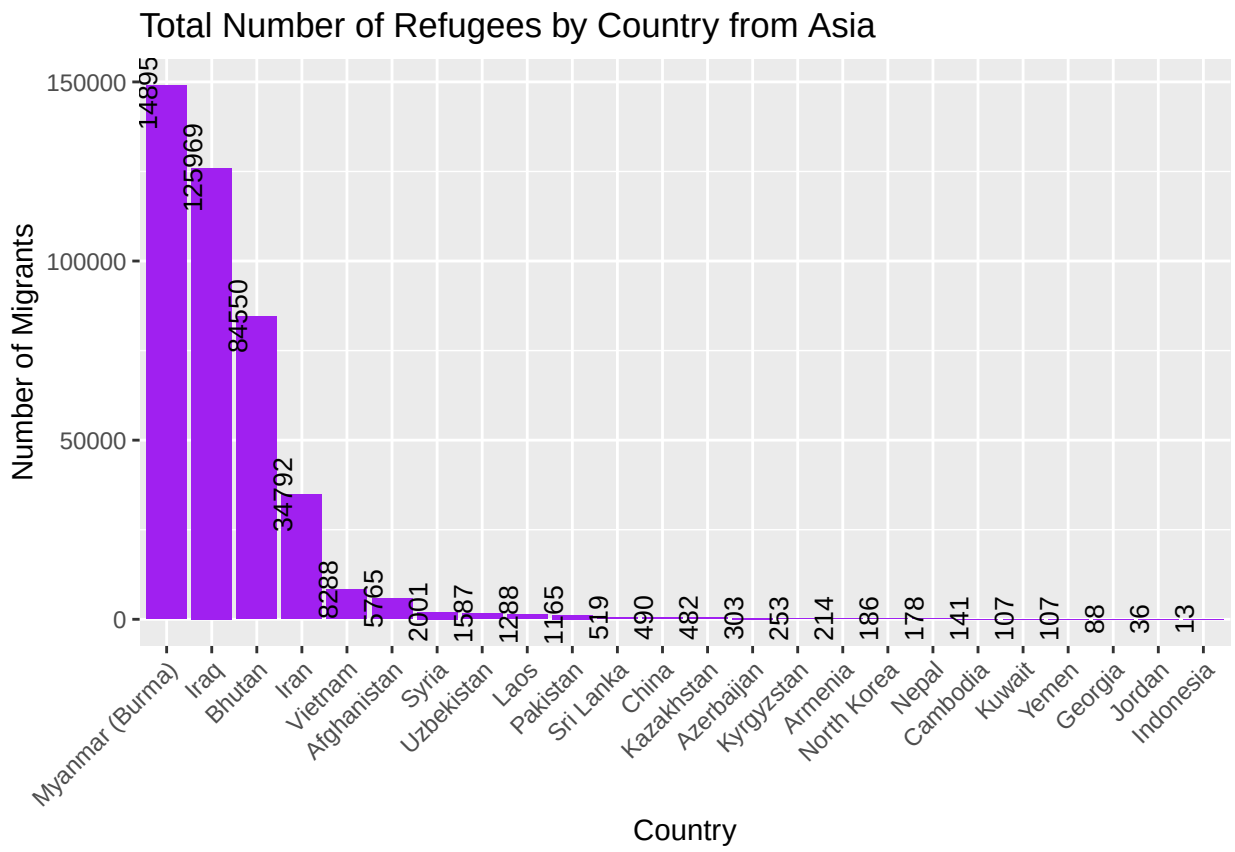
```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'Côte d’Ivoire' in 'mbcsToSbcs': dot substituted for <99>
```



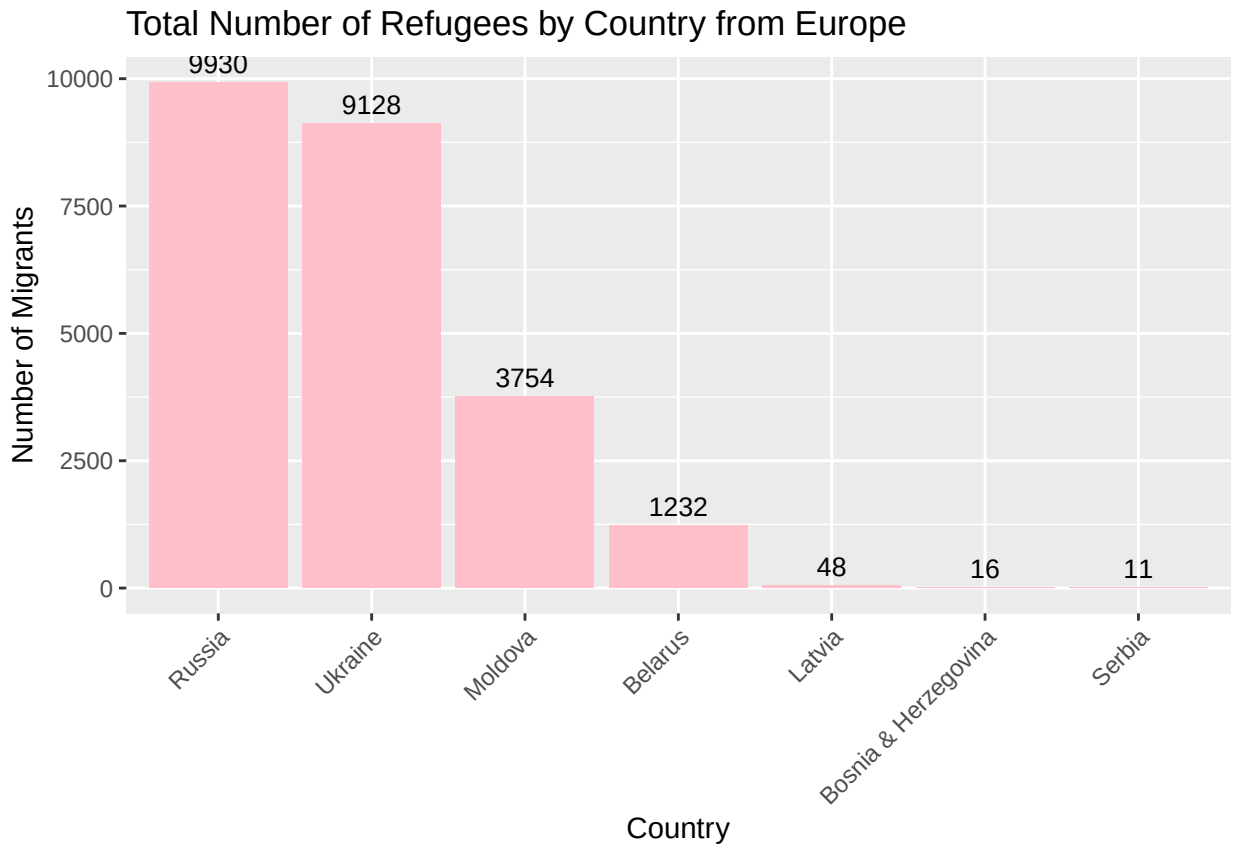
```r
americas_refugees <- refugees_clean[refugees_clean$origin_continent == "Americas", ]
americas_summary <- aggregate(number ~ origin_country, data = americas_refugees, FUN = sum)
americas_summary <- americas_summary[americas_summary$number > 1, ]

ggplot(americas_summary, aes(x = reorder(origin_country, -number), y = number)) +
  geom_bar(stat = "identity", fill = "darkgreen") +
  geom_text(aes(label = number), vjust = -0.5, size = 3.5) +
  labs(title = "Total Number of Refugees by Country from the Americas", x = "Country", y = "Number of Mi
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



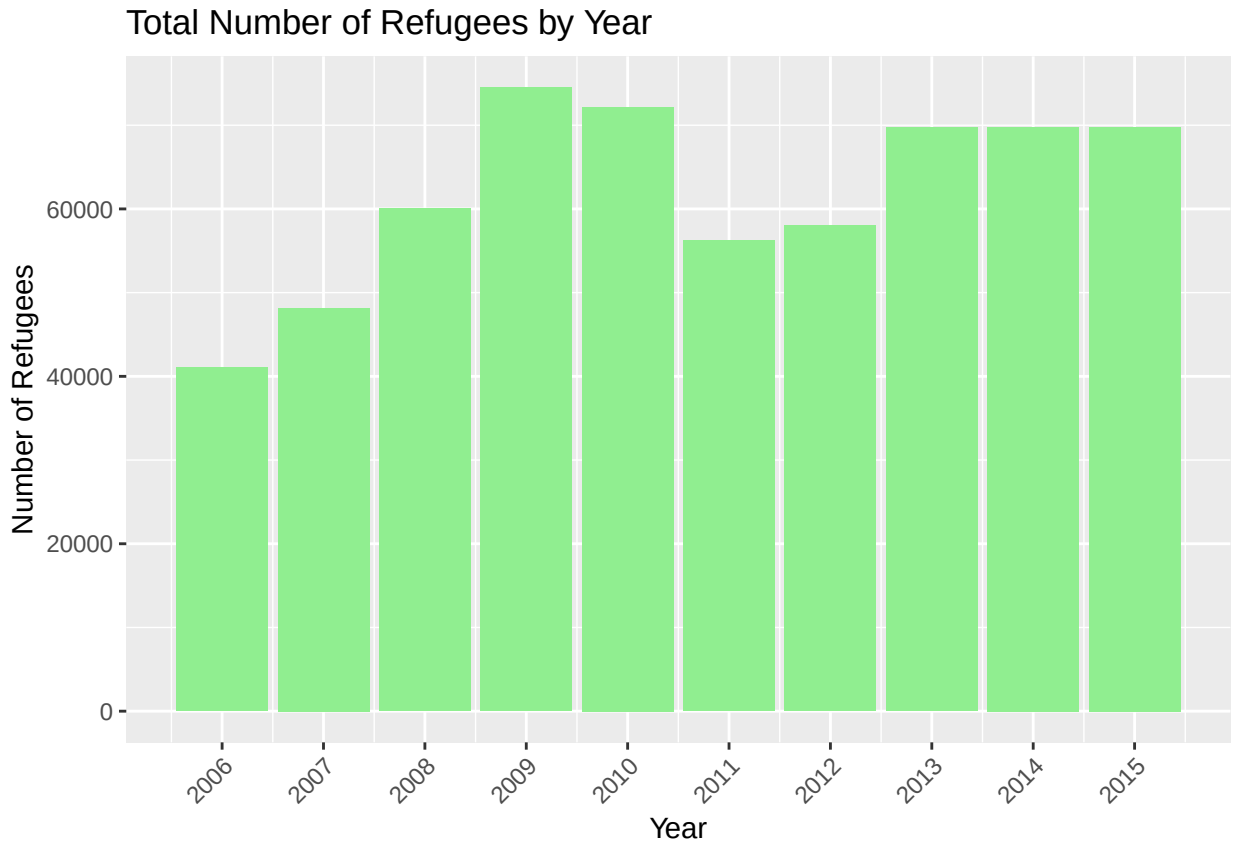
```r
asia_refugees <- refugees_clean[refugees_clean$origin_continent == "Asia", ]
asia_summary <- aggregate(number ~ origin_country, data = asia_refugees, FUN = sum)
asia_summary <- asia_summary[asia_summary$number > 1, ]

ggplot(asia_summary, aes(x = reorder(origin_country, -number), y = number)) +
  geom_bar(stat = "identity", fill = "purple") +
  geom_text(aes(label = number), vjust = -0.5, size = 3.5, angle=90) +
  labs(title = "Total Number of Refugees by Country from Asia", x = "Country", y = "Number of Migrants") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```r
europe_refugees <- refugees_clean[refugees_clean$origin_continent == "Europe", ]
europe_summary <- aggregate(number ~ origin_country, data = europe_refugees, FUN = sum)
europe_summary <- europe_summary[europe_summary$number > 1, ]

ggplot(europe_summary, aes(x = reorder(origin_country, -number), y = number)) +
  geom_bar(stat = "identity", fill = "pink") +
  geom_text(aes(label = number), vjust = -0.5, size = 3.5) +
  labs(title = "Total Number of Refugees by Country from Europe", x = "Country", y = "Number of Migrants
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



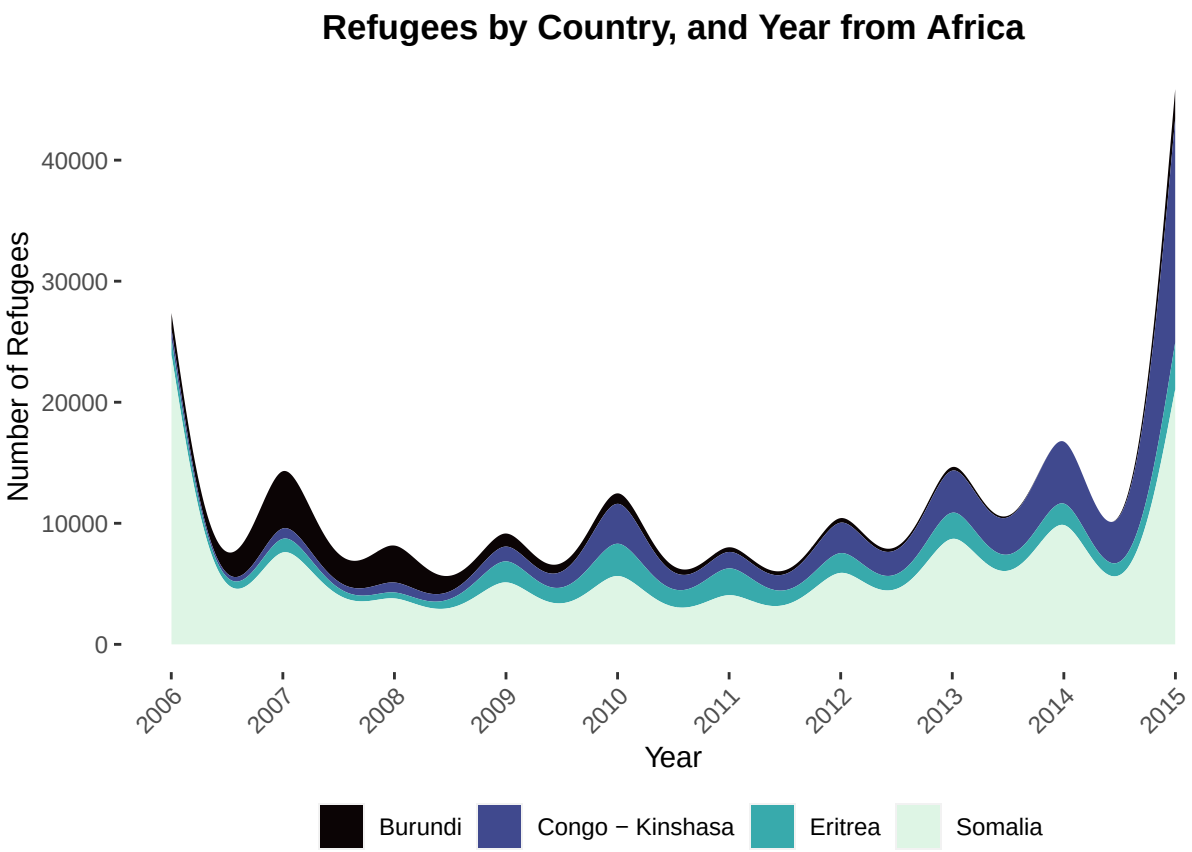
```r
# Subsetting total migrants by year to see most popular years
summary_by_year <- aggregate(number ~ year, data = refugees_clean, FUN = sum)

library(ggplot2)

ggplot(summary_by_year, aes(x = year, y = number)) +
  geom_bar(stat = "identity", fill = "lightgreen") +
  labs(title = "Total Number of Refugees by Year", x = "Year", y = "Number of Refugees") +
  scale_x_continuous(breaks = summary_by_year$year) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```
library(viridis)
```

```
## Loading required package: viridisLite
```

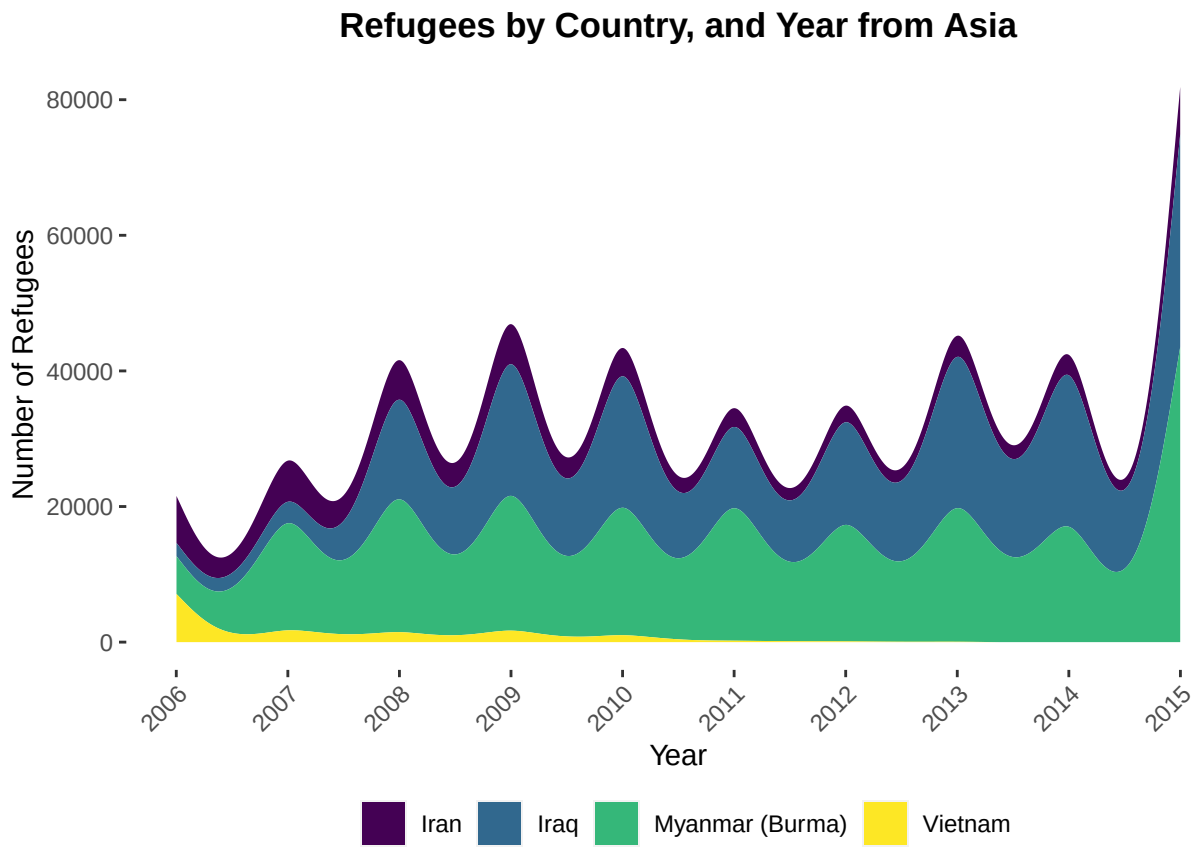
```
library(ggtext)
library(ggstream)

refugees_africa <- refugees_clean %>%
  filter(origin_continent == "Africa")

total_refugees_by_country <- refugees_africa %>%
  group_by(origin_country) %>%
  summarize(total_refugees = sum(number)) %>%
  arrange(desc(total_refugees)) %>%
  top_n(4, total_refugees)

refugees_top_4 <- refugees_africa %>%
  filter(origin_country %in% total_refugees_by_country$origin_country)

ggplot(refugees_top_4, aes(x = year, y = number, fill = origin_country)) +
  geom_stream(type = 'ridge') +
  scale_fill_viridis(discrete = TRUE, option = "mako") +
  scale_x_continuous(breaks = unique(refugees_top_4$year), labels = unique(refugees_top_4$year)) +
  labs(title = "<b><span style='text-shadow: 2px 2px gray'>Refugees by Country, and Year from Africa</sp
  theme(panel.background = element_rect(fill = "white"),
        axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_markdown(hjust = 0.5),  # use element_markdown() for ggtext formatting
        legend.position = "bottom",
        legend.box = "horizontal",
        legend.margin = margin(t = 0, r = 0, b = 0, l = 0),
        legend.spacing = unit(0.2, "cm"))
```




```
ggsave("refugees_top_4_Africa.png")
```

```
## Saving 6.5 x 4.5 in image
```

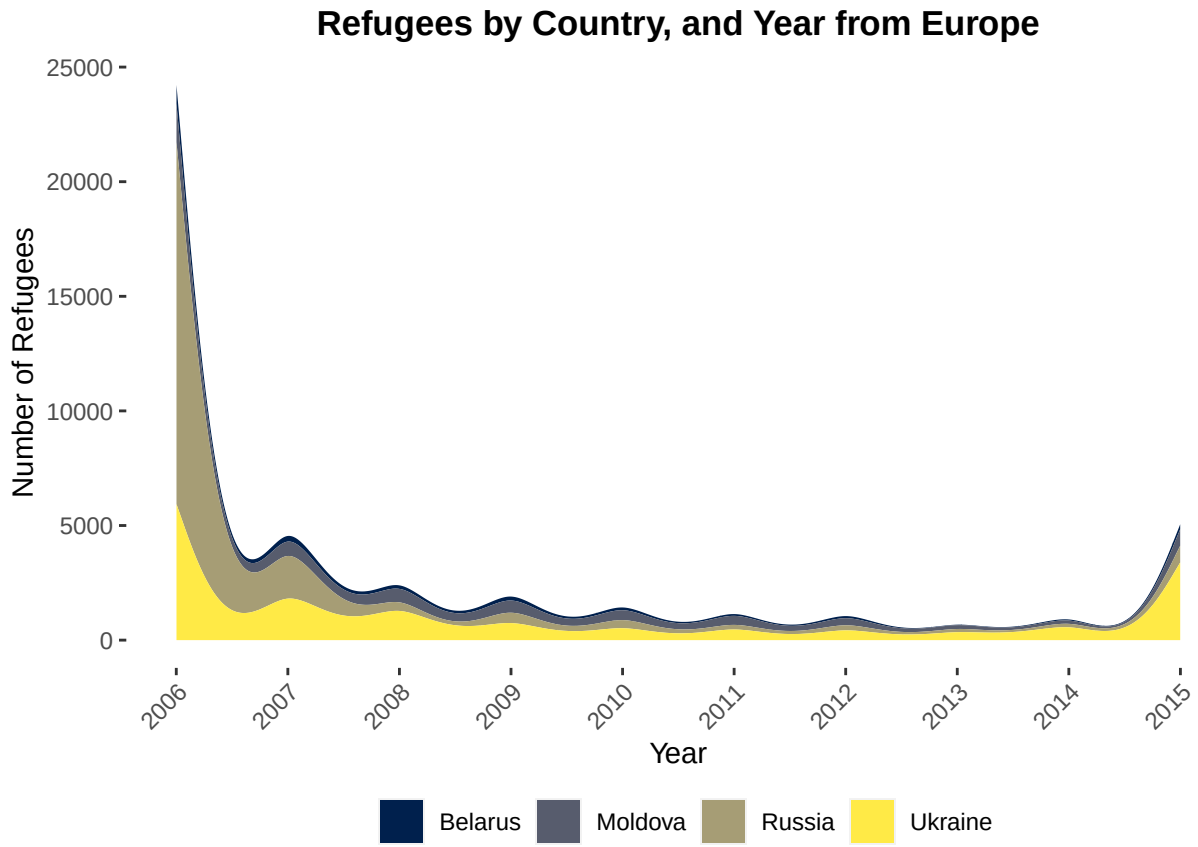
```r
refugees_asia <- refugees_clean %>%
  filter(origin_continent == "Asia")

total_refugees_by_country_asia <- refugees_asia %>%
  group_by(origin_country) %>%
  summarize(total_refugees = sum(number)) %>%
  arrange(desc(total_refugees)) %>%
  top_n(4, total_refugees)

refugees_top_4_asia <- refugees_asia %>%
  filter(origin_country %in% total_refugees_by_country_asia$origin_country)

ggplot(refugees_top_4_asia, aes(x = year, y = number, fill = origin_country)) +
  geom_stream(type = 'ridge') +
  scale_fill_viridis(discrete = TRUE, option = "viridis") +
  scale_x_continuous(breaks = unique(refugees_top_4_asia$year), labels = unique(refugees_top_4_asia$yea
  labs(title = "<b><span style='text-shadow: 2px 2px gray'>Refugees by Country, and Year from Asia</spa
  theme(panel.background = element_rect(fill = "white"),
        axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_markdown(hjust = 0.5),  # use element_markdown() for ggtext formatting
        legend.position = "bottom",
        legend.box = "horizontal",
        legend.margin = margin(t = 0, r = 0, b = 0, l = 0),
        legend.spacing = unit(0.2, "cm"))
```



```
ggsave("refugees_top_4_Asia.png")
```

```
## Saving 6.5 x 4.5 in image
```

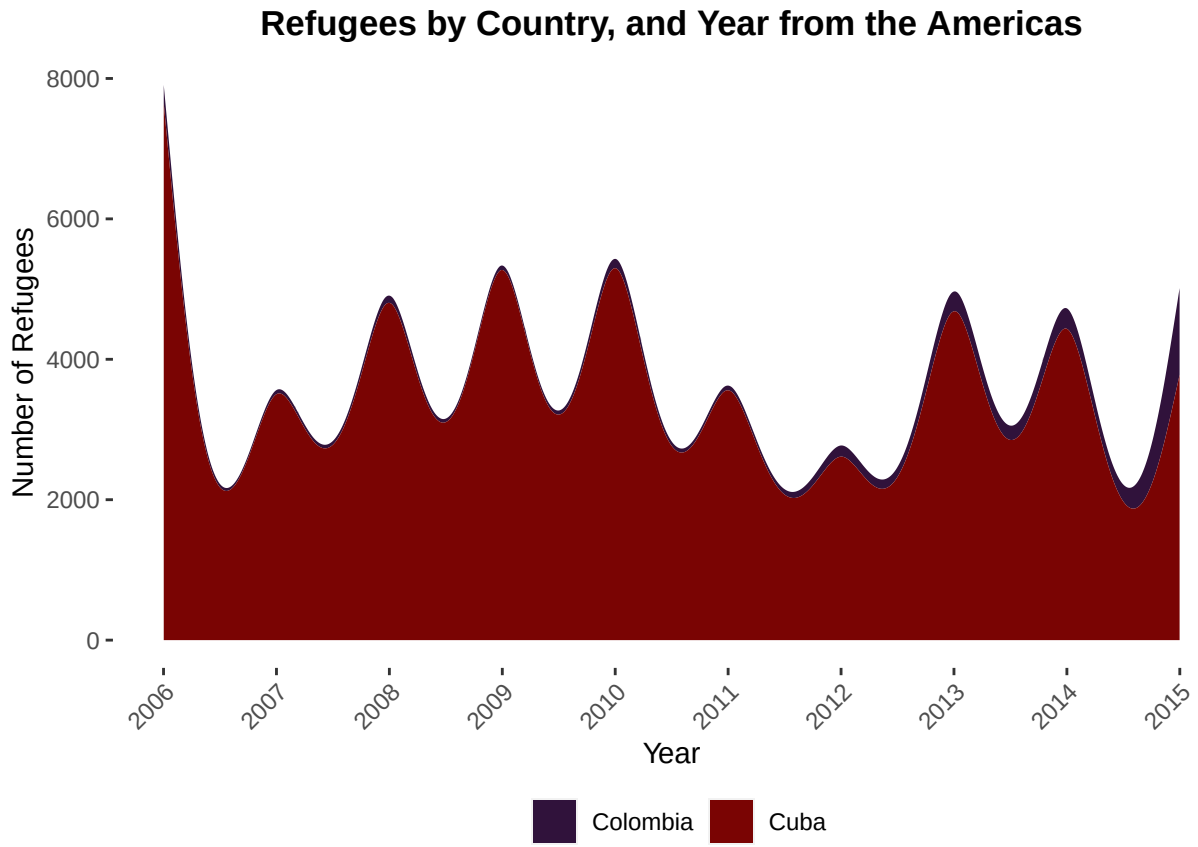
```r
refugees_europe <- refugees_clean %>%
  filter(origin_continent == "Europe")

total_refugees_by_country_europe <- refugees_europe %>%
  group_by(origin_country) %>%
  summarize(total_refugees = sum(number)) %>%
  arrange(desc(total_refugees)) %>%
  top_n(4, total_refugees)

refugees_top_4_europe <- refugees_europe %>%
  filter(origin_country %in% total_refugees_by_country_europe$origin_country)

ggplot(refugees_top_4_europe, aes(x = year, y = number, fill = origin_country)) +
  geom_stream(type = 'ridge') +
  scale_fill_viridis(discrete = TRUE, option = "cividis") +
  scale_x_continuous(breaks = unique(refugees_top_4_europe$year), labels = unique(refugees_top_4_europe$
  labs(title = "<b><span style='text-shadow: 2px 2px gray'>Refugees by Country, and Year from Europe</sp
  theme(panel.background = element_rect(fill = "white"),
        axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_markdown(hjust = 0.5),  # use element_markdown() for ggtext formatting
        legend.position = "bottom",
        legend.box = "horizontal",
        legend.margin = margin(t = 0, r = 0, b = 0, l = 0),
        legend.spacing = unit(0.2, "cm"))
```



```
ggsave("refugees_top_4_Europe.png")
```

```
## Saving 6.5 x 4.5 in image
```

```r
refugees_amer <- refugees_clean %>%
  filter(origin_continent == "Americas")

total_refugees_by_country_amer <- refugees_amer %>%
  group_by(origin_country) %>%
  summarize(total_refugees = sum(number)) %>%
  arrange(desc(total_refugees)) %>%
  top_n(4, total_refugees)

refugees_top_4_amer <- refugees_amer %>%
  filter(origin_country %in% total_refugees_by_country_amer$origin_country)

ggplot(refugees_top_4_amer, aes(x = year, y = number, fill = origin_country)) +
  geom_stream(type = 'ridge') +
  scale_fill_viridis(discrete = TRUE, option = "turbo") +
  scale_x_continuous(breaks = unique(refugees_top_4_amer$year), labels = unique(refugees_top_4_amer$yea
  labs(title = "<b><span style='text-shadow: 2px 2px gray'>Refugees by Country, and Year from the Ameri
  theme(panel.background = element_rect(fill = "white"),
        axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_markdown(hjust = 0.5),  # use element_markdown() for ggtext formatting
        legend.position = "bottom",
        legend.box = "horizontal",
        legend.margin = margin(t = 0, r = 0, b = 0, l = 0),
        legend.spacing = unit(0.2, "cm"))
```



```
ggsave("refugees_top_4_Americas.png")
```

```
## Saving 6.5 x 4.5 in image
```