Final

Team 9

2024-02-24

Introduction

The United States has always been home to countless individuals who have chosen to leave their countries of birth. America has served as a beacon of hope for individuals uprooted by tragedy. Our study unravels the delicate fabric of asylum, documenting the silent journeys across continents as we delve into the patterns and experiences of the refugees seeking solace on American beaches and lands between 2006 and 2015.

```r
#install.packages("maps")
#install.packages("countrycode")

library(tidyverse)
library(sf)
library(maps)
library(plotly)
library(gapminder)
library(forcats)
library(scales)
library(countrycode)
library(ggrepel)


#install.packages("stringdist")
# library(stringdist)

refugee <- read.csv("data/A2_refugee_status.csv")
world_map <- read_sf("data/ne_110m_admin_0_countries/ne_110m_admin_0_countries.shp")
world_sans_antarctica <- world_map %>%
  filter(ISO_A3 != "ATA")
```

Data Cleaning

```r
refugee[refugee == 'D' | refugee == '-' | refugee == 'X'] <- 0
continents <- c("Asia", "Africa", "Europe", "North America", "South America")

refugee_continent <- refugee[refugee$Continent.Country.of.Nationality %in% continents, ]

colnames(refugee_continent)[
  colnames(refugee_continent) == "Continent.Country.of.Nationality"] <- "continent"
```

```
refugee_continent_tidy <- pivot_longer(data = refugee_continent,
                                        cols = -c(continent),
                                        names_to = "year",
                                        values_to = "count")
# remove the x in the beginning of year column
refugee_continent_tidy$year <- sub("^X", "", refugee_continent_tidy$year)
refugee_continent_tidy$count <- gsub(",", "", refugee_continent_tidy$count)

refugee_continent_tidy$year <- as.numeric(refugee_continent_tidy$year)
refugee_continent_tidy$count <- as.numeric(refugee_continent_tidy$count)
```

Refugee per Continent over Time

```
# Calculate the total count of each continent
continent_totals <- aggregate(count ~ continent, data = refugee_continent_tidy, sum)

# Reorder the levels of the continent factor variable based on total count
refugee_continent_tidy$continent <- factor(
  refugee_continent_tidy$continent,
  levels = continent_totals$continent[order(continent_totals$count,
                                            decreasing = TRUE)])

# Define colors for each continent
continent_colors <- c(
  "Africa" = "#FFA07A",      # Light Salmon
  "Asia" = "#4682B4",        # Steel Blue
  "Europe" = "#2E8B57",      # Sea Green
  "North America" = "#FF6347",  # Tomato
  "Oceania" = "#6A5ACD",     # Slate Blue
  "South America" = "#FFD700"   # Gold
)

# Plot with reordered color legend and custom line colors
chart1 <- ggplot(data = refugee_continent_tidy,
           mapping = aes(
             x = year,
             y = count,
             color = continent
           )) +
  geom_line(size = 1) +
  scale_x_continuous(
    breaks = seq(2006, 2015, by = 2),
    labels = function(x) ifelse(
      x == 2006, as.character(x), paste0("'", substr(x, 3, 4)))) +
  labs(
    x = "Year",
    y = "People",
    color = "Continent"
  ) +
  scale_color_manual(values = continent_colors) +  # Assigning custom colors
  theme(
```
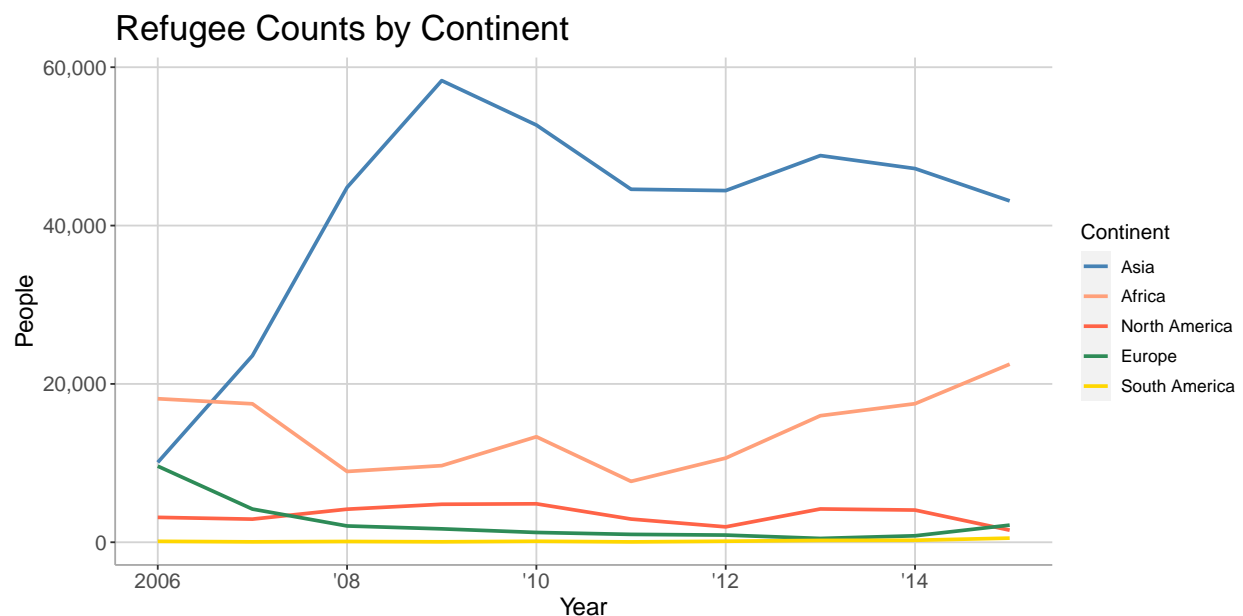
```
    plot.title = element_text(size = 20),
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.title = element_text(size = 12),
    legend.text = element_text(size = 10),
    panel.background = element_rect(fill = "white"),
    panel.grid.major = element_line(color = "lightgray"),
    panel.grid.minor = element_blank(),
    axis.line = element_line(color = "darkgray")
  ) +
  ggtitle("Refugee Counts by Continent")+
  scale_y_continuous(labels = comma)

ggsave(chart1, filename = "output/chart1.pdf", width = 10, height = 5)
ggsave(chart1, filename = "output/chart1.png", width = 10, height = 5)


chart1
```


The influx of refugees from 2006 to 2015 tells a story of perseverance, global politics, and humanitarian response. As the world experienced the highest levels of displacement since World War II, the United States became a reflection of these global shifts, with its demographic fabric reflecting the ebb and flow of international crises.

The timeline chart provides valuable insight into the trends of refugees accepted in the US. This tool is excellent for understanding patterns and changes over time. The audience can examine the migration of individuals who immigrated to America from various continents between 2006 and 2015. Notably, Asia exhibits distinct behaviour compared to other continents. It's crucial to recognize that Asia, with its significant population, holds particular importance in this analysis.

From this graph, we can glean two key insights: 1) There has been a steady rise in the number of refugees since 2006, with Asia experiencing its most significant peak in 2009, marking a remarkable 58% increase.

2) Since 2011, the influx of refugees from Africa has continued to rise steadily without any signs of decline. What could be driving these trends in these two regions of the world?

Asia's 2009 peak likely stemmed from conflicts and disasters in countries like Afghanistan, Iraq, and Myanmar. Africa's steady refugee rise since 2011 ties to civil wars, instability, and crises in nations such as Somalia, South Sudan, and the DRC. These trends reflect complex geopolitical factors forcing millions to flee, and the US has responded by providing refuge amid humanitarian emergencies.

Asia Refugee Distribution

```r
valid_countries <- c(
    "Afghanistan", "Angola", "Armenia", "Azerbaijan", "Belarus", "Bhutan",
    "Bosnia-Herzegovina", "Burma", "Burundi", "Cambodia", "Cameroon",
    "Central African Republic", "Chad", "China, People's Republic",
    "Colombia", "Congo, Democratic Republic", "Congo, Republic",
    "Cote d'Ivoire", "Cuba", "Egypt", "Equatorial Guinea", "Eritrea",
    "Ethiopia", "Gambia", "Georgia", "Haiti", "Honduras", "Indonesia",
    "Iran", "Iraq", "Jordan", "Kazakhstan", "Kenya", "Korea, North",
    "Kuwait", "Kyrgyzstan", "Laos", "Latvia", "Liberia", "Mauritania",
    "Moldova", "Nepal", "Nigeria", "Pakistan", "Russia", "Rwanda",
    "Serbia", "Sierra Leone", "Somalia", "South Sudan", "Sri Lanka",
    "Sudan", "Syria", "Togo", "Uganda", "Ukraine", "Uzbekistan",
    "Vietnam", "Yemen", "Zimbabwe"
)

# Delete rows that are not country names
refugee_country <- subset(refugee, Continent.Country.of.Nationality %in% valid_countries)
colnames(refugee_country)[
  colnames(refugee_country) == "Continent.Country.of.Nationality"] <- "country"


# longer the country data
refugee_country_tidy <- pivot_longer(data = refugee_country,
                                     cols = -c(country),
                                     names_to = "year",
                                     values_to = "count")

# year and count to numeric
refugee_country_tidy$year <- sub("^X", "", refugee_country_tidy$year)
refugee_country_tidy$count <- gsub(",", "", refugee_country_tidy$count)

refugee_country_tidy$year <- as.numeric(refugee_country_tidy$year)
refugee_country_tidy$count <- as.numeric(refugee_country_tidy$count)


refugee_country_average <- refugee_country_tidy %>%
  group_by(country) %>%
  summarise(average = mean(count))

refugee_country_average$iso_code <- countrycode(
  refugee_country_average$country, "country.name", "iso3c")

asia_map <- merge(world_sans_antarctica,
```

4

```
                     refugee_country_average,
                     by.x = "ISO_A3", by.y = "iso_code",
                     all.y = TRUE) %>%
    filter(CONTINENT == "Asia")


asia_map <- asia_map %>%
    mutate(should_be_labeled =
        ifelse(country %in% c("Burma", "Iraq", "Iran"), TRUE, FALSE))



label_asia <- filter(asia_map, should_be_labeled == TRUE)



# Get centroids of the MULTIPOLYGON geometries
centroids <- st_centroid(label_asia)

# Extract x and y coordinates of the centroids
centroid_coords <- st_coordinates(centroids)

# Create label data frame
label_data <- data.frame(
  country = label_asia$country,
  x = centroid_coords[, "X"],
  y = centroid_coords[, "Y"]
)

# Create the plot
chart2 <- ggplot() +
  # Add the base map for Asia
  geom_sf(data = filter(world_sans_antarctica, CONTINENT == "Asia"),
          fill = "white",
          color = "darkgray",
          size = 0.1) +
  # Add the map for refugee counts
  geom_sf(data = asia_map,
          aes(fill = average),
          size = 0.1,
          color = "darkgray",
          alpha = 0.8) +
  geom_text(
          data = label_data,
          aes(label = country, x = x, y = y),
          size = 3,
          color = "white",
          fill = "white"
  )+
  # Define the fill color scale
  scale_fill_gradient(name = "Annual Refugee",
                      low = "lightgray", high = "#2B5C8D",
                      trans = "log10",
                      labels = scales::comma) +
  # Add titles and labels
  labs(title = "Refugees in Asia",
```

```
      subtitle = "Burma, Iraq, and Iran have the highest number of refugees in Asia. ",
      fill = "Average Refugee") +
  theme_minimal() +  # Example theme; you can adjust as per your preference
  theme(plot.title = element_text(size = 20, color = "#333333"),
        plot.subtitle = element_text(size = 10, color = "#333333"),
        axis.text=element_blank(),
        axis.ticks=element_blank(),
        axis.title = element_blank(),
        legend.title = element_text(size = 10, color = "#333333"),
        legend.text = element_text(size = 8, color = "#333333"),
        plot.caption = element_text(size = 10, color = "gray"),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(color = "darkgray", size = 0.8, fill = NA)) +
  coord_sf()

ggsave(chart2, filename = "output/chart2.pdf", width = 8, height = 5)
ggsave(chart2, filename = "output/chart2.png", width = 8, height = 5)


chart2
```
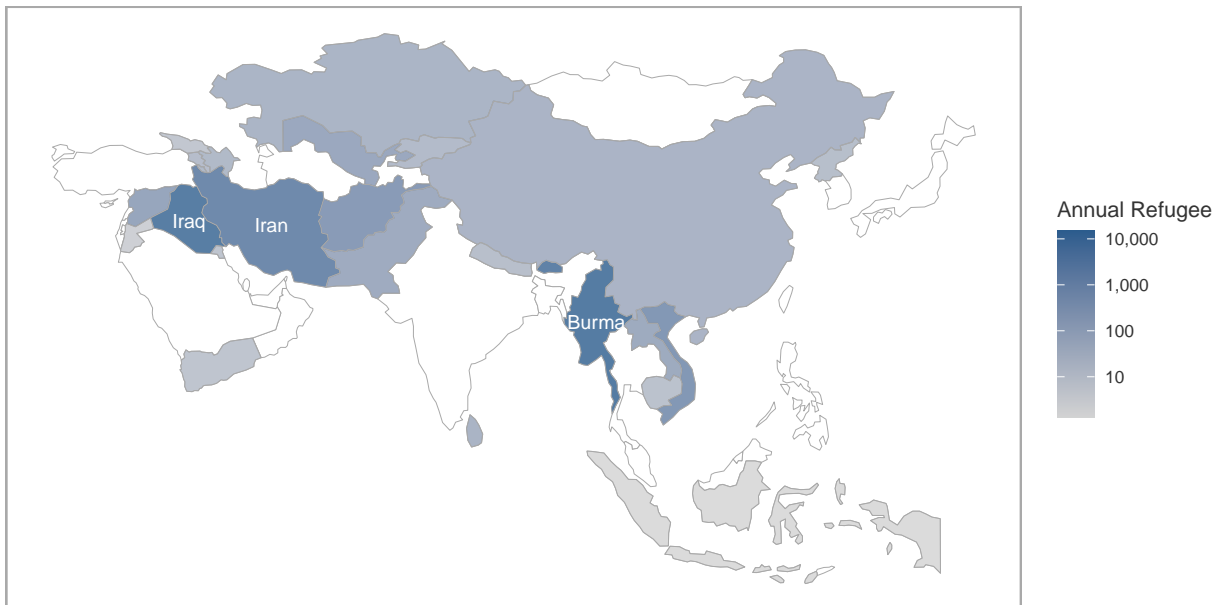
Refugees in Asia

Burma, Iraq, and Iran have the highest number of refugees in Asia.

Our focus is primarily on Asia, as it accounts for the largest refugee population and overall population worldwide. Iraq, Myanmar, and Iran stand out as the top three countries seeking refugee status in the USA. The visual narrative begins in 2008, a time of relative openness, and progresses through to a more restrictive 2015, capturing the lives behind the numbers as refugees settled across the states, with the US becoming a new home for many.

Additionally, Asia has not been immune to the global recession that has impacted the Western world since 2008. In 2009, nations like Vietnam faced deficits in their current account balances, while Korea grappled with escalating unemployment rates due to the economic crisis.

The use of heat maps helps the audience easily visualize and interpret the distribution of refugee populations from various countries over time. The colour gradient allows for clear representation of population volumes, making patterns and trends more readily apparent. Heat maps are an effective way to convey complex data in an easily digestible visual format.

Zooming in on Asia, the continent that, according to Pew Research, has contributed more than half of the refugees to the United States since 1980. Our visualization dissects the top five countries within Asia, examining the factors leading to their prominence in this period. The story we tell here is nuanced and complex, portraying the confluence of geopolitical forces - from entrenched conflicts to economic disruptions - that compelled people from Iraq, Myanmar, and other nations to seek sanctuary in distant lands.

Africa Refugee Timeline

```r
refugee_country_tidy$iso_code <- countrycode(
  refugee_country_tidy$country, "country.name", "iso3c")
african_countries <- merge(
  world_sans_antarctica,
  refugee_country_tidy,
  by.x = "ISO_A3",
  by.y = "iso_code",
  all.y = TRUE)

african_countries <- filter(african_countries, CONTINENT == "Africa")
highlight_countries <- c("Congo, Democratic Republic", "Somalia")

# Assuming 'african_countries' is your dataset

# Add a new column 'color' to indicate whether the country is in highlight_countries
african_countries <- african_countries %>%
  mutate(color =
           ifelse(country %in% highlight_countries,
                  country, "Other African Countries"))
```

```r
refugee_continent_tidy$continent <- factor(
  refugee_continent_tidy$continent,
  levels = continent_totals$continent[order(continent_totals$count, decreasing = TRUE)])

african_countries$color <- factor(
  african_countries$color,
  levels = c("Somalia", "Congo, Democratic Republic", "Other African Countries"))

# Define custom color palette
custom_colors <- c(
  "Somalia" = "#8B0000",  # Dark red
  "Congo, Democratic Republic" = "#006400",  # Dark green
  "Other African Countries" = "lightgray"  # Light grey
)

# Create the plot
```
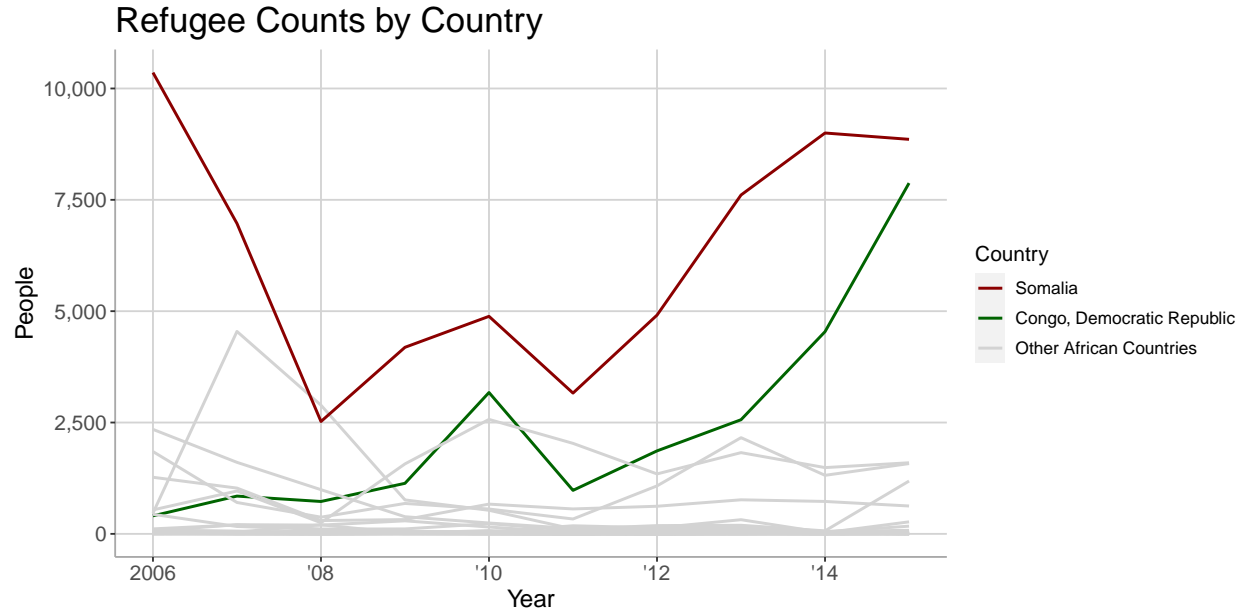
```r
chart3 <- ggplot(data = african_countries,
       aes(x = year, y = count, color = color, group = country)) +
  geom_line(size = 0.8) +
  scale_color_manual(values = custom_colors) +
  #theme_minimal() +
  labs(
    x = "Year",
    y = "People",
    color = "Country"
  ) +
  scale_color_manual(values = custom_colors) +  # Assigning custom colors
  theme(
    plot.title = element_text(size = 20),
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.title = element_text(size = 12),
    legend.text = element_text(size = 10),
    panel.background = element_rect(fill = "white"),
    panel.grid.major = element_line(color = "lightgray"),
    panel.grid.minor = element_blank(),
    axis.line = element_line(color = "darkgray")
  ) +
  ggtitle("Refugee Counts by Country")+
  scale_y_continuous(labels = comma) +
  scale_x_continuous(
    breaks = seq(1990, 2021, by = 2),
    labels = function(x) ifelse(
      x %in% c(2006), as.character(x), paste0("'", substr(x, 3, 4))))) 


ggsave(chart3, filename = "output/chart3.pdf", width = 10, height = 5)
ggsave(chart3, filename = "output/chart3.png", width = 10, height = 5)

chart3
```

Refugee Counts by Country

The use of the timeline is crucial in our analysis, as we seek to compare refugee trends across Africa from 2006 to 2015, grouping countries into three categories: the Democratic Republic of Congo, Somalia, and other countries. A line chart proves to be the most suitable visualization option for this purpose.

Focusing specifically on the continent, we home in on the years 2010, as this marks a significant turning point where Congo and Somalia exhibit similar behaviours in their refugee trends. It is during this time that both countries experience their first important peak in refugee numbers, with the population surge continuing unabated since 2011. In Somalia, this surge is exacerbated by the dire effects of the largest drought in history, compelling people to flee their homes in search of necessities like food, water, and assistance (Hujale, 2022). Meanwhile, the Democratic Republic of Congo faces its own challenges, including a failed coup attempt and attacks in Kinshasa, amidst decades of armed conflict, human rights abuses, and gender-based violence.

As we delve deeper into the data, the narratives of Congo and Somalia, alongside other nations, emerge as pivotal in understanding the broader wave of migration from Africa. These stories of conflict and hope encapsulate the struggles and resilience of African refugees.

Through this exploration, we aim to foster a deeper understanding of the impact these refugees have had on America and to honour the courage and perseverance embodied in their journeys. Utilizing a line chart, we will effectively illustrate the trends and patterns across the three African groups from 2006 to 2015, enabling us to compare their refugee flows.