INF1340H Final Assignment
Visualization of United Nations International Migrant Stock Datasets
Cynthia Yip
17 December, 2021

Introduction

United Nations' dataset 'Trends in International Migrant Stock: The 2015 Revision' provides information on world population, international migrant stock, percentage of migrant stock in total population, the gender distribution of migrants, and refugee stock from 1990 to 2015. This paper uses tools in exploratory data analysis like scatter plots, line plots, box plots, and violin plots to identify trends, patterns, and relationships in this international migrant stock dataset. An overview of the worldwide migration data shows that while there is an upward trend of the worldwide international migrant stock, the migrant stock as a percentage of the world population has remained relatively stable over this span of 25 years. By further breaking down the dataset by regions, major areas, top destinations of international migrants, and gender, the report examines whether the international migrant stock data across various countries have also been stable, or there are changes in migration trends and migrant status characteristics over time.

Methods and Results

Re-cleaning

The resulting datasets from the previous tidy data assignment include 1) datasets with international migrant stock as the observation unit: 'untable_year_region' and 'untable_year_country' and 2) datasets with the annual rate of change of migrant stock as the observation unit: 'untable_range_region' and 'untable_year_country'. However, since many columns in these datasets have 'Python object' data type, additional data cleaning is performed to transform them into the appropriate data types (i.e. 'Python object', 'int' and 'float') for further processing in exploratory data analysis. Spaces in numerical numbers (e.g., in the columns 'Population (thousands)' and 'International Migrant Stock') are removed. Data entries with missing values denoted as '..' in the datasets are replaced with NaN. These cleaned datasets are exported to new CSV files: 'untable_year_region.csv', 'untable_year_country.csv', 'untable_range_region.csv' and 'untable_range_country.csv'.

Exploratory Data Analysis

This exploratory data analysis report starts by examining 1) the world's international migrant stock. Then, it moves on to look into 2) the various regions (i.e. 'developed regions', 'developing regions', 'less developed regions excluding least developed countries', and 'least developed countries'); 3) the major areas (i.e., 'Europe', 'NorthernAmerica', 'Oceania', 'Japan', 'Africa', 'Asia', 'LatinAmericaandtheCaribbean'); and 4) the top destinations of international migrants.

1) World

To only examine the data on the total world population, I filter the 'untable_year_region.csv' with the conditions 1) 'Destination' as 'WORLD' and 2) 'Gender' as 'Both'.

Figure 1

International Migrant Stock vs International Migrant Stock Percentage of World Population) (1990 – 2015)

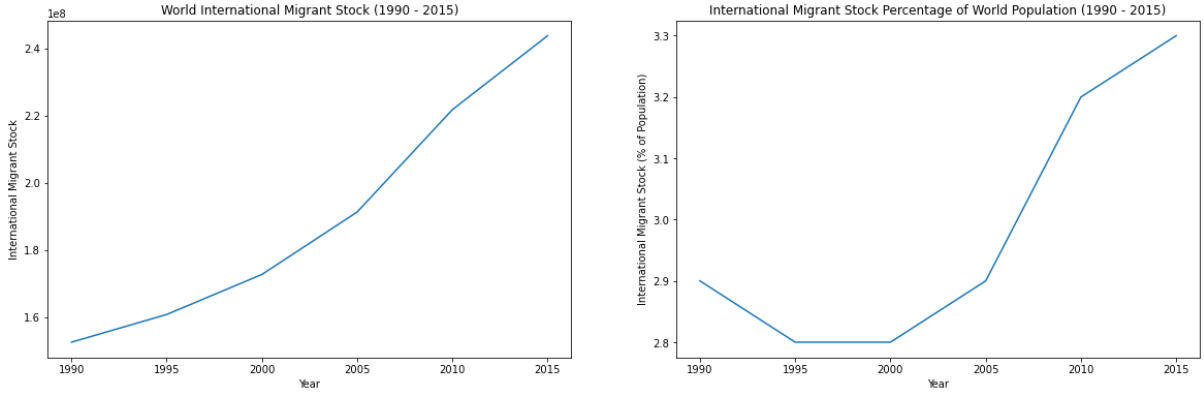


Figure 1 shows the line plots of the worldwide international migrant stock and the international migrant stock as a percentage of the world population from 1990 to 2015. The absolute number of international migrant stock has increased by 200% over the years. The global migrant stock as a percentage of the population also shows an upward trend. It has been mostly steady from 1990 to 2005 with a more significant increase from 2005 to 2010. Overall, the share of international migrant stock in the world's population has been relatively steady.

2) Regions

To examine data on regions, I filter the 'untable_year_region.csv' with the conditions 1) 'Destination' as 'Developed regions', 'developing regions', 'less developed regions excluding least developed countries' and 'least developed countries'. Although 'developing regions' includes 'less developed regions excluding least developed countries' and 'least developed countries', these two categories are included for comparison purposes.

Figure 2

Distribution of International Migrant Stock among Destinations

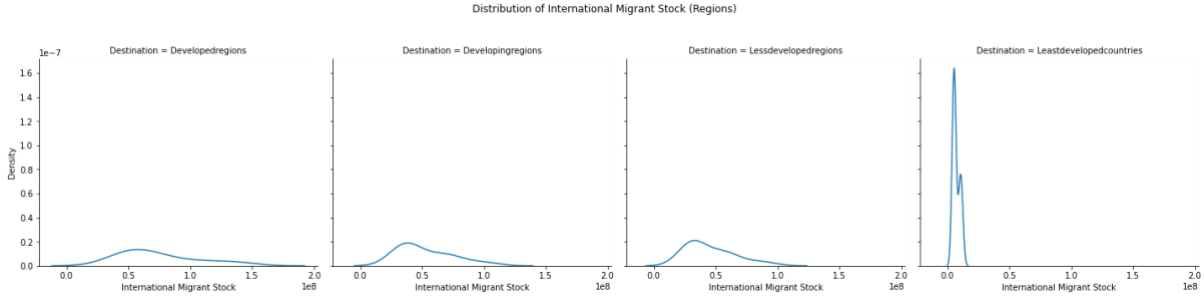


Figure 3

Distribution of International Migrant Stock (% of Population) among Destinations

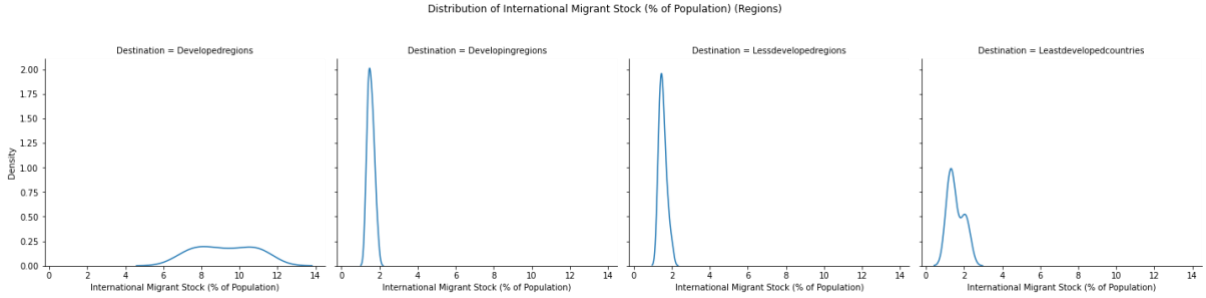
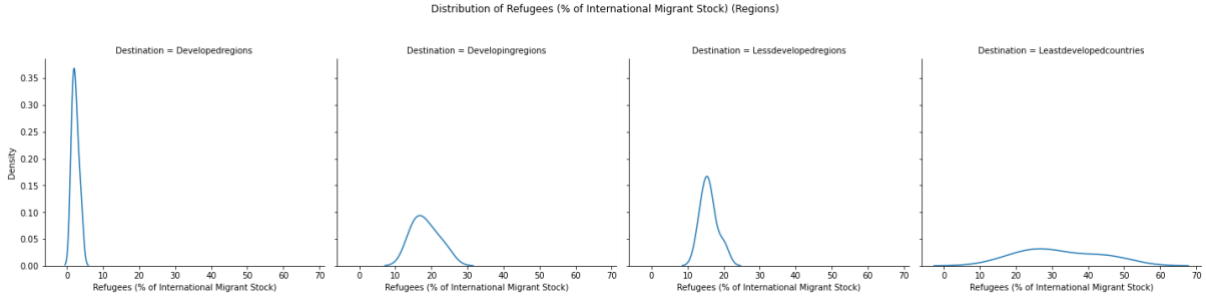


Figure 4

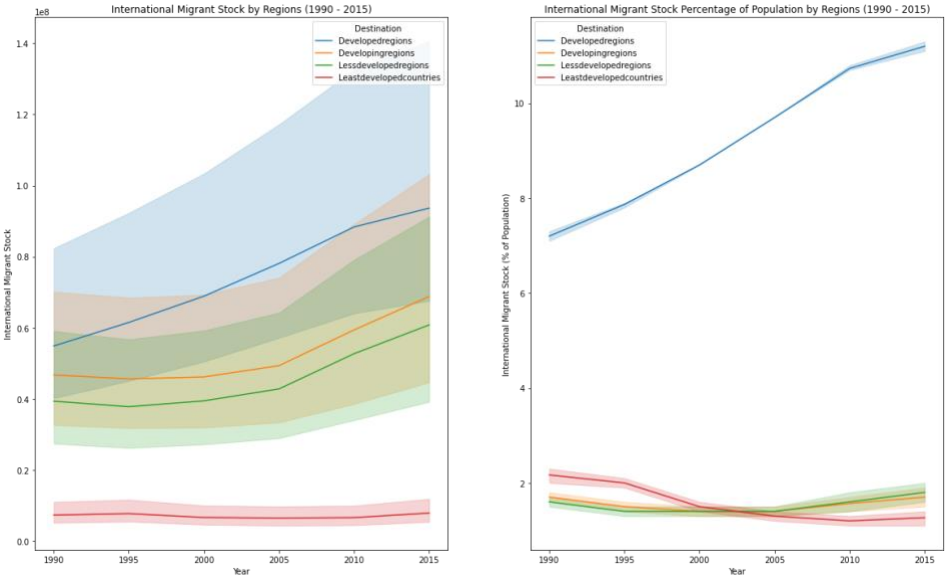
Distribution of Refugees (% of International Migrant Stock) among Destinations



Figures 2 - 4 show the distribution of the different regions against the variables international migrant stock, international migrant stock as a percentage of the population, and refugees as a percentage of the international migrant stock. These graphs help identify the patterns and differences among the different regions. Figure 2 shows a similar distribution of the absolute number of international migrant stock among the regions with the various development stages. So international migrant stock is not a good variable in distinguishing the regions. Figure 3 and 4, on the other hand, shows more differences between developed countries and developing countries. Figure 3, showing the international migrant stock as a percentage of the population, indicates a more significant proportion of their population is international migrants (ranging from 4 - 14%).Contrary to developed regions, developing regions' proportion of international migrants in their total population is much smaller (ranging from around 1 - 2%). Less developed regions composed of developing regions without the least developed countries also show a similar distribution. Figure 4, showing refugees as a percentage of international migrant stock, presents a distinct difference between developed and developing regions. The developed regions contain fewer refugees in their international migrant stock than developing regions.

Figure 5

International Migrant Stock by Regions vs International Migrant Stock Percentage of population by Regions (1990 – 2015)



Similar to figure 1, figure 5 contains two line plots showing the international migrant stock and international migrant stock as a percentage of the population of different regions from

1990 to 2015. Having hue as the destinations allows comparison between the various regions. Figure 5 shows that developed regions have a steady increase from 1990 to 2010 in terms of the absolute number of international migrant stock and the international migrant stock as a percentage of the population. Developing regions show an upward trend in the number of international migrant stock, but its percentage in their total population remains relatively steady.

3) Major Areas

According to United Nations' definition, more developed regions comprise Europe, Northern America, Australia/New Zealand (Oceania) and Japan. And less developed regions comprise all regions of Africa, Asia (except Japan), Latin America and the Caribbean plus Melanesia, Micronesia and Polynesia. Now that we have a better understanding of the distribution and characteristic of international migrant stock among regions with different development stages, we will now break it down to look at the data by the major areas. I filter the 'untable_region_year.csv' with the country codes of the major areas (i.e., Europe, Northern America, Australia and New Zealand (Oceania), Africa, Asia, and Latin America, and the Caribbean. As Japan is excluded from Asia in United Nations' definition, data on Japan is obtained by filtering 'untable_country_year.csv'.

Similar to the previous section on the regions, we first plot the distribution of the dataset against different variables using FacetGrid.

Figure 6

Distribution of International Migrant Stock (Major Areas)

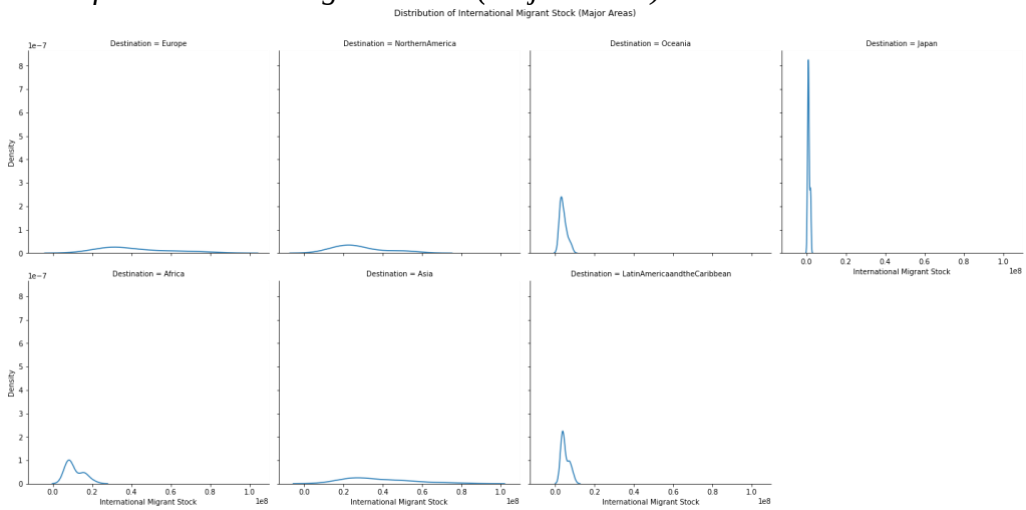


Figure 7

Distribution of International Migrant Stock (% of Population) (Major Areas)

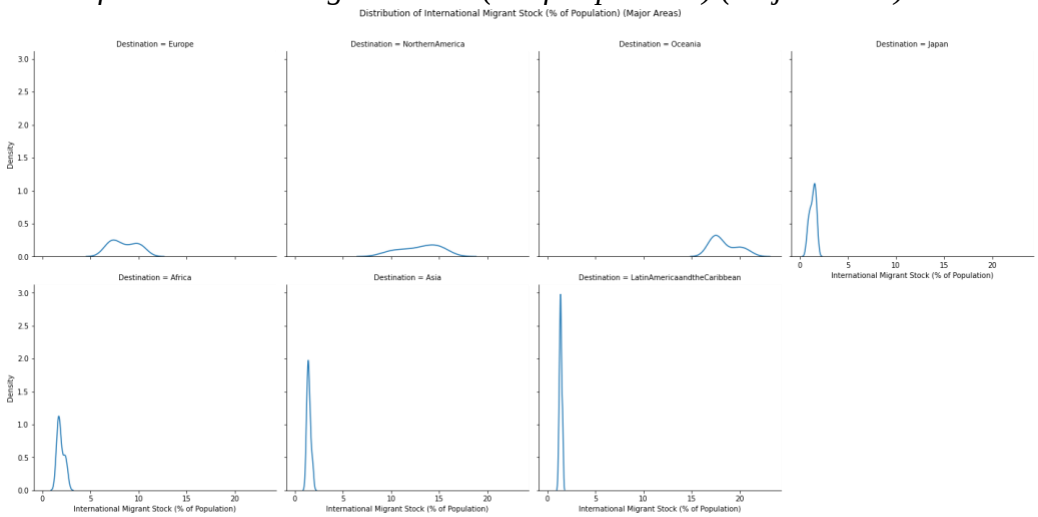
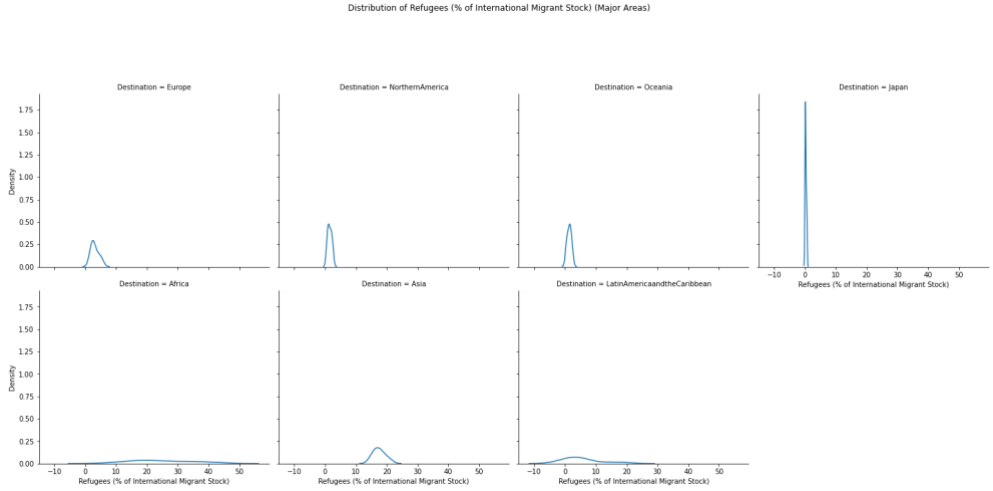


Figure 8

Distribution of Refugees (% of International Migrant Stock) (Major Areas)



Figures 6-8 show the distribution of international migrant stock, international migrant stock as a percentage of the population, and refugees as a percentage of international migrant stock for the seven major areas. In terms of international migrant stock, Europe, Northern America, and Asia have a similar distribution. In terms of the distribution of international migrant stock as a percentage of the population, all four developed major areas present very different compositions of migrants in their population. Japan has the least proportion of migrants in their population, followed by Europe and Northern America. Oceania has a very high concentration of migrants in their population (over 15%). On the other hand, developing major areas show a more similar pattern in migrant composition, all having around 0-3% of their population as international migrants. As for the percentage of refugees in the international migrant stock, developed major areas show a similar pattern, with Japan having close to 0% of their migrants as refugees. Developing regions have a wider distribution of the percentage of refugees in their population.

Figure 9

Population by Major Areas vs International Migrant Stock by Major Areas (1990 – 2015)

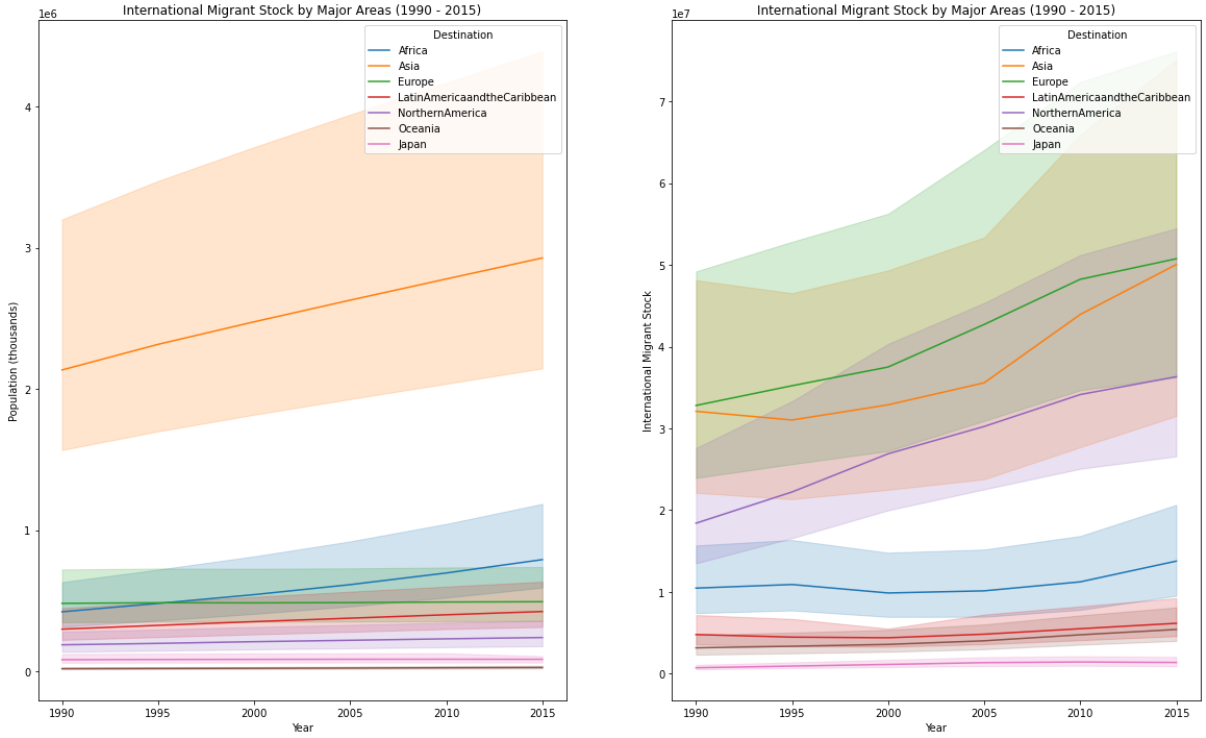


Figure 9 shows major areas' changes in population and international migrant stock from 1990 to 2015. From these two line plots, we can see that for the developed areas like Europe and North America, their population has remained largely unchanged from 1990 to 2015, even with the steady increase in international migrant stock. This shows that international migrants largely offset their declining population. Developing countries like Asia have a steady increase in their total population and international migrant stock, which means that the region maintains a positive change in its population regardless of the number of migrants. The same goes for Africa. Its population has increased even though its migrant stock has been relatively steady over the years.

Figure 10

International Migrant Stock by Major Areas vs International Migrant Stock (% of Population) by Major Areas (1990 – 2015)

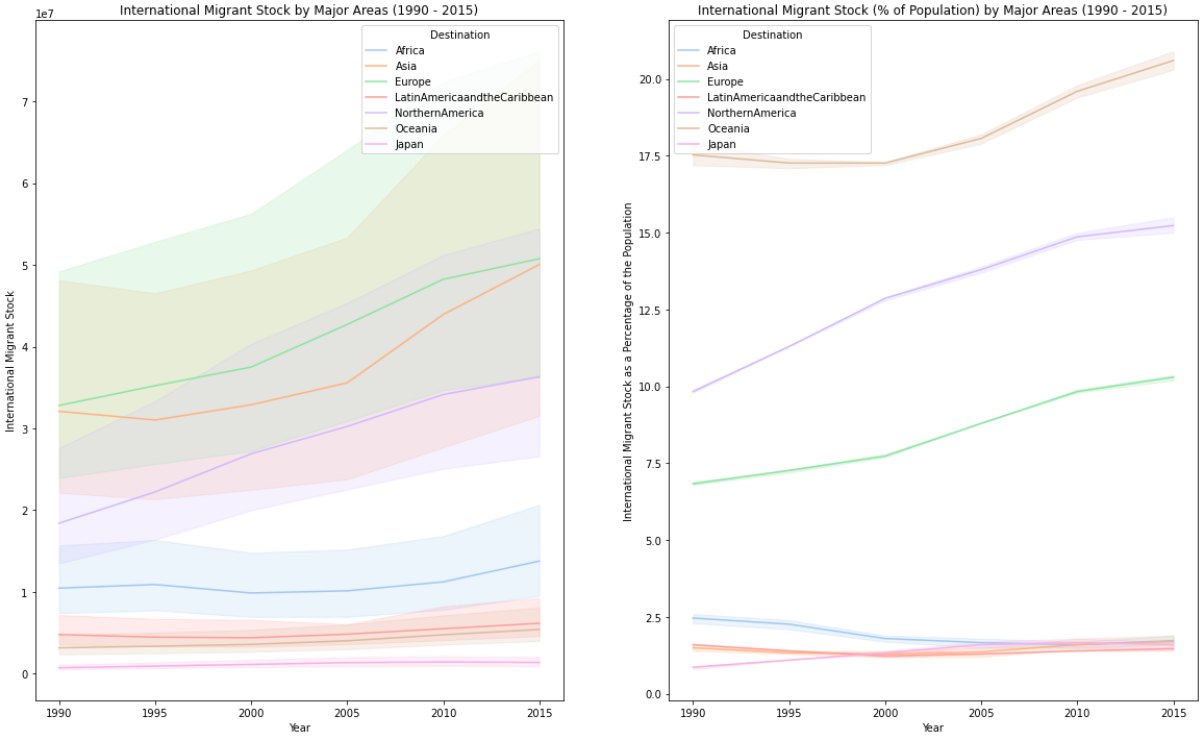


Figure 10 shows the line plots of the international migrant stock and international migrant stock as a percentage of the population of different regions from 1990 to 2015. Overall, there has been an increase in the number of international migrants in all regions. Europe and Asia are the two areas with the most significant growth. In particular, Asia has had a sharp increase in its migrant stock since 2005. But when looking at the graph that measures the international migrant stock as a proportion of their population. It shows that Oceania has the greatest growth of the percentage of international migrants in its population. It also has the highest proportion of migrants among all regions, with migrants representing over 20% of its population in 2015. Both Northern America and Europe show an upward trend, but the increase slowed down in 2000. Even though Asia has grown substantially in terms of the absolute number of international migrant stock, the share of international migrant stock in its total population is still relatively small and steady throughout 1990 to 2015.

Female Migrants

To study the data on the changes of proportion of female migrants among international migrants, I filter the previous major area DataFrame (‘target_majory’) with the condition of ‘Gender’ as ‘Female’.

Figure 12

Boxplot and Violinplot of Female Migrant Stock (% of Population) by Major Areas (1990 - 2015)

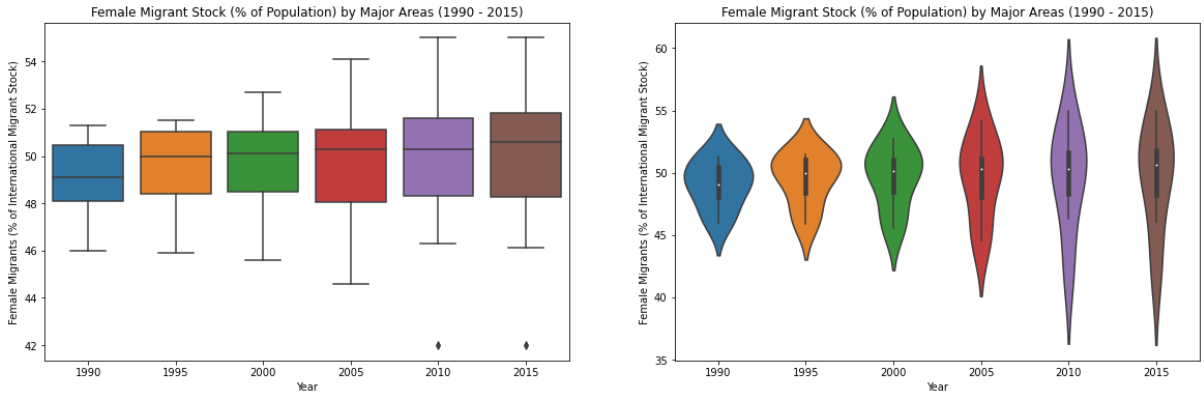


Figure 12 presents a boxplot and a violin plot of the percentage of female migrant stock in the total migrant stock from 1990 to 2015. The share of female migrant stock has slightly increased, with the medium going from around 49% in 1990 to over 50% in 2015. However, the boxplot and violin plot show that the percentage of female migrants among all major areas is relatively concentrated from 1990 to 2000. Starting from 2005, the percentage of female migrants is more spread out among the areas with more extreme values in both ends, particularly in the lower end with outliers of having only 42% of the international migrant stock as female.

Figure 13

Female Migrant Stock (% of Population) by Major Areas (1990 - 2015)

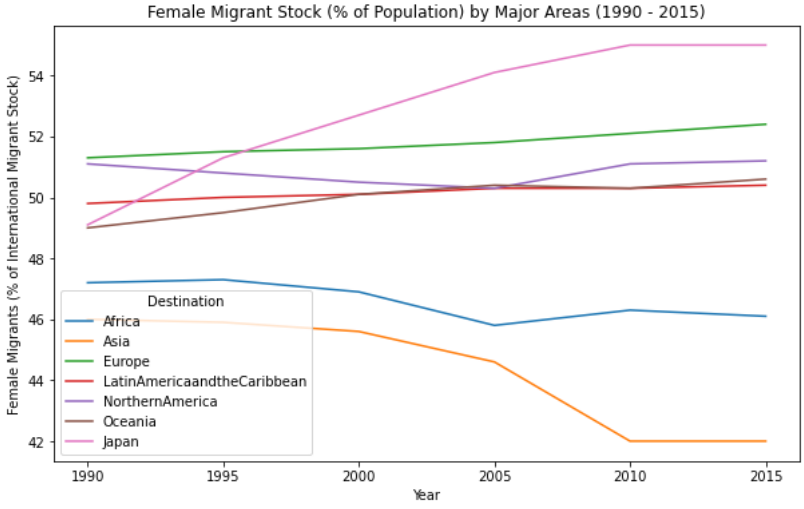


Figure 13 presents the same data as figure 12 with information on major areas. There are certain parts of the world where male migrants outnumber female migrants and vice versa. Japan, Europe, and Oceania have over half of their migrant population as female. They also have an increasing trend since 1990. Northern America shows a relatively steady gender distribution and fair representation in their migrant population. On the other hand, developing regions show a decrease in the percentage of female migrants from 1990 to 2015. Particularly Asia that has had a sharp decline in female representation since 2000.

4) Top Destinations of International Migrants (countries)

Lastly, to get a better understanding on the top destination countries of International Migrant, I filter the 'untable_country_year.csv' with the conditions of countries with the top number of migrants and 'Gender' as 'Both'.

Figure 13

International Migrant Stock by Top Destinations (1990 - 2015)

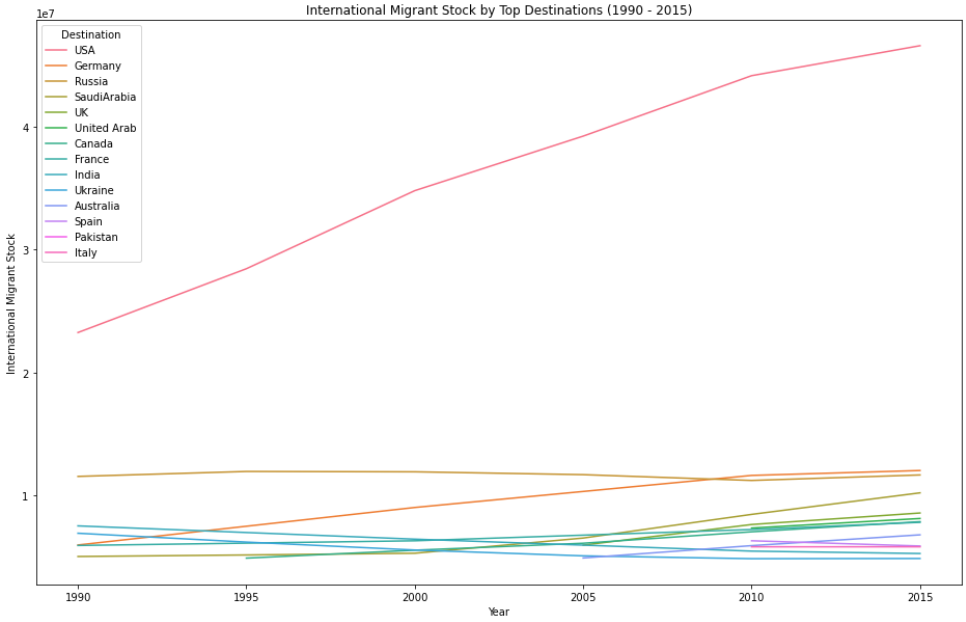


Figure 13 shows the top migration destinations from 1990 to 2015. Among these 14 destinations, the United States of America has always been the top destination. Its number also increases exponentially, widening its gap with the second top destination. It outnumbers the second top destination by almost 40 millions in 2015. European countries like Germany are gaining popularity, growing from the fifth in 1990 to the second most popular destination in 2015. Some not-so-traditional destinations have gained popularity. For example, some southern European countries like Spain and Italy also start to appear in the top list from 2010. Destinations in the Middle East like Saudi Arabia have also become increasingly popular since 2000. It grew to become the fourth most popular destination in 2015. Still, the majority of these countries are in developed regions.

Figure 14

Distribution of International Migrant Stock by Top Destinations (1990 - 2015)

Figure 14 visualizes the distribution of the international migrant stock of these top destinations throughout the years with the violin plot. The overall shape and distribution are similar, and the medium is stable. But there are more and more outliers towards the higher end throughout the years, which should be the United States of America data as shown in figure 13.

Figure 15

Population vs International Migration Stock vs % of Population by Top Destinations

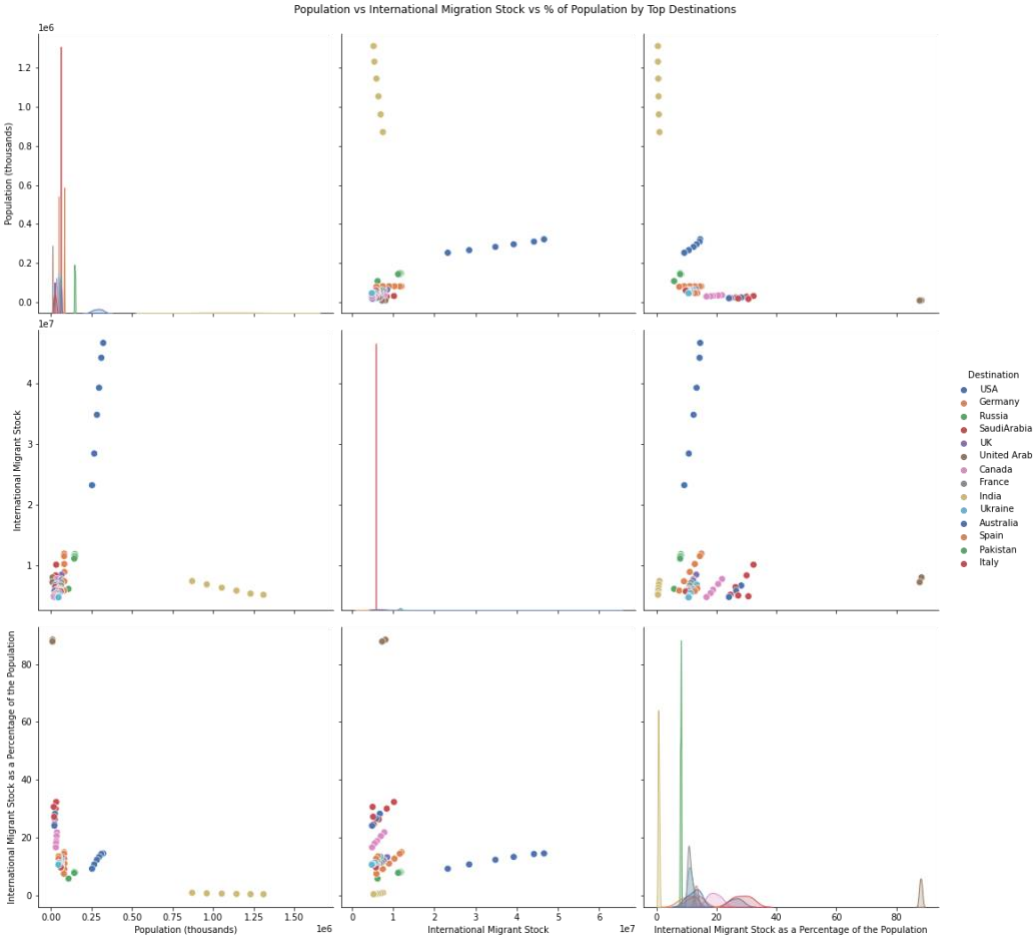


Figure 15 is the pair plot of population, international migrant stock, and international migrant stock as a percentage of the population of the top destinations. Setting the hue as the destinations shows that these destinations share different characteristics. The United States of America is characterized by its large international migrant stock, but it shares a similar proportion of migrants in its total population as other European countries. Canada and Australia are similar in terms of their population and the percentage of international migrant stock in their population. The United Arab Emirates has the highest proportion of migrants in their population. India has the highest population but relatively lower international migrant stock compared to the other destinations.

Implications

From the above exploratory data analysis, we can conclude that even though the proportion of international migrant stock in the world population remains relatively stable from 1990 to 2015, the landscape of international migrants among developed and developing regions, various major areas, and top destinations has changed.

Migrants' considerations when choosing a destination have changed. Figures 5, 8, and 9 show the changing trends of international migrants in developed and developing regions and various major areas. While the United States of America remains the most popular destination, other top countries have shifted their ranking in 25 years. Since all these destinations have different characteristics, these changes show the changing needs of the migrants themselves and changing international migration policies and openness towards migration of the various countries. As a result, some countries like Germany grow faster than others in obtaining international migrant stock. In addition, apart from the more traditionally popular destinations, cities in southern Europe and oil-producing nations in the Middle East are gaining popularity. Developing countries are also seeing a more significant number of growth of international migrants. However, developed countries still dominate the top of the migration destination list.

International migrants are more critical to certain countries than others. Developed countries are characterized by their high proportion of international migrants in their total population. In comparing their population growth and international migrant stock growth in figure 9, we can see that international migrants are more important to developed countries in maintaining their population than developing countries. For example, in Australia and New Zealand, their population is relatively small but with a very high proportion of migrants. Thus, these countries need to continue retaining and attracting more migrants to balance their age, and the labor force needs to maintain their economic growth.

The analysis also presents two major challenges of international migration. The first one is an imbalance in supporting refugees globally. The refugee population is more concentrated in developing countries. A refugee crisis is characterized by citizens being forcibly displaced from their home countries because of violence, war, or natural disasters (United Nations n.d.). As refugee crises usually happen in politically less stable regions, the higher proportion of refugees in developing countries may also be because they are closer to their home countries. They are also usually less-skilled workers who may be less attractive to developed countries.

Throughout the years, the changing gender composition among the various areas shows characteristics of the areas and presents potential challenges. Developed countries like Northern America and Europe have been relatively attractive to female migrants. Developing

countries like Asia are relatively more attractive to male migrants. In particular, Asia experiences a sharp decrease in the percentage of female migrants in their population. In addition, as presented in the finding, this gap is widening. This gap presents different challenges for both developed and developing countries with the underrepresentation of a particular gender in their international migrant stock.

Conclusion

Performing exploratory data analysis of the United Nations' international migrant stock dataset provides a better understanding of the changes in international migrant stock around the globe from 1990 to 2015. Given the various social, economic, political changes in different countries, there is a changing landscape of various regions and the top destinations, showing a shift of reasons for people to migrate. International migrants are more vital to certain countries than others, so maintaining their attractiveness to migrants is very important in sustaining their development. Finally, the differences in refugee composition and gender distribution in developed and developing regions remain some of the biggest challenges in international migration.

Reference

McKinney, W., & others. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (Vol. 445, pp. 51–56).

Anon. n.d. 'Choosing Color Palettes — Seaborn 0.11.2 Documentation'. Retrieved 17 December 2021a (https://seaborn.pydata.org/tutorial/color_palettes.html).

Anon. n.d. 'Python - How to Add a Title to Seaborn Facet Plot'. *Stack Overflow*. Retrieved 17 December 2021b (https://stackoverflow.com/questions/29813694/how-to-add-a-title-to-seaborn-facet-plot).

Anon. n.d. 'Seaborn FacetGrid Plot Two Different y Axis next to Eachother - Pretag'. Retrieved 17 December 2021d (https://pretagteam.com/question/seaborn-facetgrid-plot-two-different-y-axis-next-to-eachother).

T, Bex. 2020. 'Master a Third of Seaborn: Statistical Plotting with Relplot()'. *Medium*. Retrieved 17 December 2021 (https://towardsdatascience.com/master-a-third-of-seaborn-statistical-plotting-with-relplot-df8642718f0f).

United Nations. n.d. 'Refugees | United Nations'. Retrieved 17 December 2021c (https://www.un.org/en/global-issues/refugees).