UN International Migrant Stock Data 2015 Revision: Data Visualization and Interpretation Report

Yifei Chen (Yvette 1002271187)
December 15, 2022
Code: https://colab.research.google.com/drive/1Yj1ETij_7XCfNphJzr40WQbeWpAPeugi?usp=sharing

Overview

The previous report focused on cleaning the United Nations dataset "Trends in International Migrant Stock: The 2015 Revision", which is an important foundation for this visualization interpretation report. The aim of this report is to use visualizations to explore trends in international migration at different levels, and answer the following questions:

1. Which continents are the most popular destinations for international migrants?
2. How does international migration differ by sex?
3. What are the overall trends in international migration?

Summary of the Visualizations Created

Table 1:International Migrant Stock	2.1 Multiple Bar Chart: Total Number of Migrants in the World by Sex 1.1 Multiple Bar Chart: Total Number of Migration by Continents 3.1 Multiple Bar Chart: Total Number of Migration by Development 2.3 Line Charts: Migration Trend by Continents, world, and Development (Small Multiple)
Table 2: Total Population	2.2 Line Chart: Total Population trend by Sex 3.2 Line Chart: Total Population trend by Region
Table 3: International Migrant Stock as a Percentage of the Total Population	2.4 Violin Plot: Migrant Percentage, by Sex
Table 4: Female Migrants as a Percentage of the International Migrant Stock	2.5 Line Chart: Female Migration Percentage by Continent (Small Multiple) 2.6 Line Chart: Female Migration Percentage in the World 2.7 Line Chart: Female Migration Percentage by Continent adjusted Y-axis (Small Multiple)
Table 5: Annual Rate of Change of the Migrant Stock	2.8 Line Chart: Annual Rate of Change in migrant stock by Sex
Table 6: Estimated Refugee Stock	3.3 Violin Plot vs Box plot: Estimated Refugee Stock by development (Small Multiple) 3.4 Line Chart: Estimated Refugee Stock

	3.5 Line Chart:Refugee Stock as a Percentage in International Migration

Table of Contents

Summary 1

Methods 2

1. Popular Continents as Migration Destinations 3

2. Understanding International Migration by Sex 3

3. Overall Trends in International Migration 10

    3.1 Refugee 12

4. Applying Visualization Principles 14

Summary

1. Europe, Asia, and Northern America are the most popular continents for international migration.

2. On a global level, there are more male migrants than female migrants throughout the studied period. But this report concludes that sex imbalance in international migration does not exist because some continents/regions have more females than males and vice versa. There are also more males in the total world population than females.
    a. On a continental level, Europe has more female migrants than males; Oceania, Latin America, and North America nearly balanced male and female migrant stock; Asia and Africa align with the global trend.
    b. From a development perspective, developed regions have more female migrants and developing regions have more male migrants.

3. On a global level, the total number of international migrants consistently increase from 1990 to 2015. The same trend applies to the regional level for developed regions. In the developing regions, while there is an overall increase in the international migrant stock, there is a slight drop from around 1995 to 2000.

4. On a global level, 2015 has a slightly larger number in refugee stock compared to 1990. However, the decreasing percentage of refugees in total international migration suggests that in recent years, more people migrate internationally for different causes other than seeking asylums.

Methods

The dataset was previously cleaned so in order to create visualizations, not a lot of further actions were required to get the dataset ready. The typical first step was to think about the purpose of the visualization, what does the data represent? What are the x and y variables? What type of chart is most suitable for the data?

Once some initial planning was done, I used the loc function to select specific rows within the cleaned dataset and create new data frames. This is useful when creating visualizations that only use data from certain areas, e.g. world data, continental data, or regional data. Additionally, some variables in the tidied tables require minor updates such as grouping or pivoting, resetting the index, and changing the datatypes before they become ready for visualizations.

Process Examples:

1. To create Figure 2.1 Multiple Bar Chart, rows with data from the whole world were selected and made into a new data frame. The purpose was to show how the international migrant stock changes over time so “Year” was set as the X variable and “international migrant stock” was set as the Y variable.

2. To create Figure 2.3, a small multiple of line charts by continent, I created a new data frame for each continent with a total of 6 data frames created. Each consists of one continent’s data from 1990 to 2015. Then I used matplotlib to create subplots so all the line charts will display in one figure. The purpose of this small multiple was to compare international migrant stock trends across continents so for each chart, the Y variable is “international migrant stock”, X is “Year”, and hue is “Sex” (allow us to see the difference between male and female).

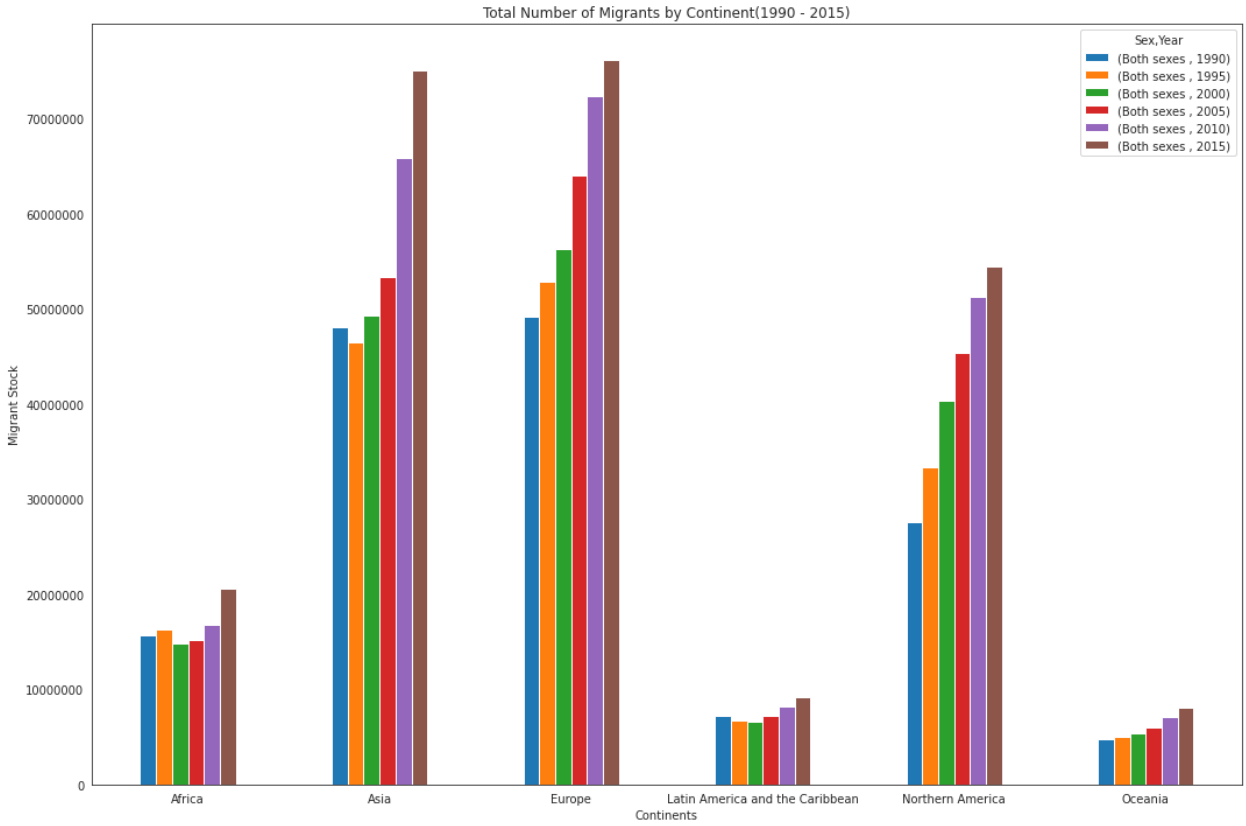
3. To create Figure 2.3, Violin plot, I applied the Seaborn function and set the kind to “violin”. The purpose of the visualization is to see the difference in the percentage of international migration stock between developed and developing regions. Thus, the X variable is “Area” (to show developed and developing regions), the Y variable is the percentage of international migration in the total population, and the hue is Sex (the plot splits by sex). I added a title for the plot using the .set(title= “ ”) function. A similar process was used to create box plots in this report.

1. Popular Continents as Migration Destinations

Research question: Which Continents are the most popular destinations for international migrants?

Figure 1.1 shows that Europe is the most popular destination, followed by Asia, then Northern America. Based on the chart, Europe and Asia seem almost equally popular but Asia experienced a decline in international migration in the mid-1990s, while Europe has had a steady increase in international migration across the years.

Figure 1.1 Number of International Migrants by Continent, 1990-2015



The decline in migration stock around 1995 also appears in the developing country chart (Figure 3.1). A hypothesis to make is that the decrease in migration in developing countries around 1995 is related to the decrease in migration in Asia because there are many major developing countries in Asia with large populations.

2. Understanding International Migration by Sex

Research question: How does international migration differ by sex?

Figure 2.1 allows us to zoom out and look at the total number of international migrants across the world from 1990 to 2015, we see a consistently increasing trend. It also shows that there are slightly more male than female international migrants in every time period. This suggests a sex imbalance in international migration, with more males than females migrating internationally.

Figure 2.1 The total number of International Migrants in the World by Sex, 1990 - 2015

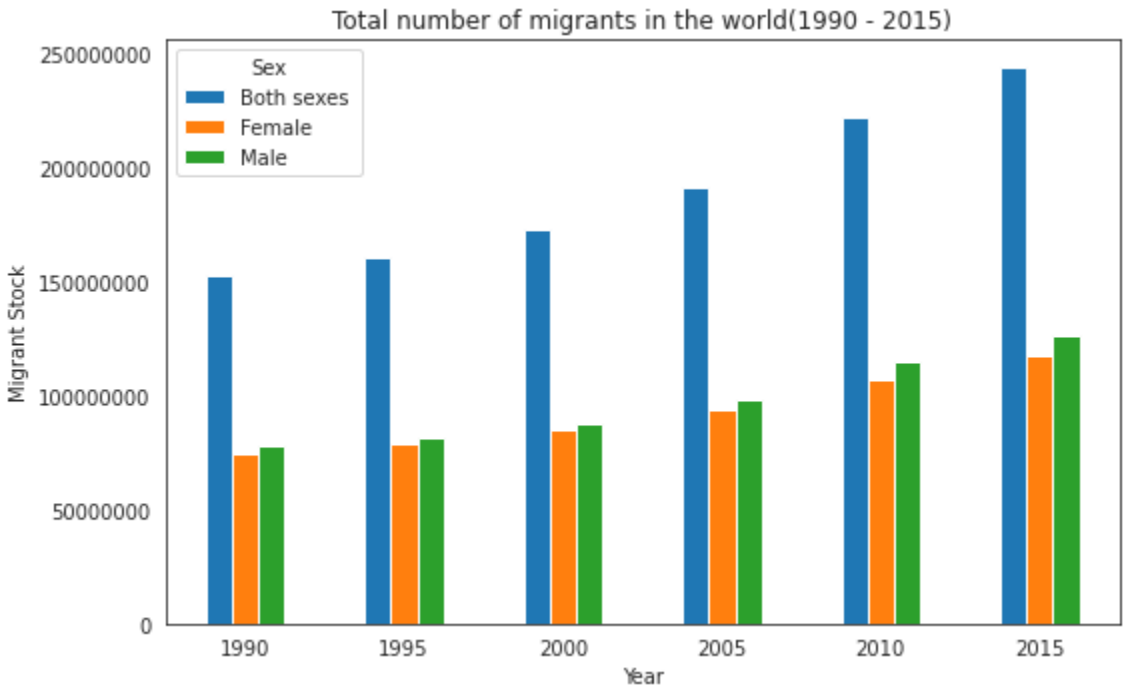
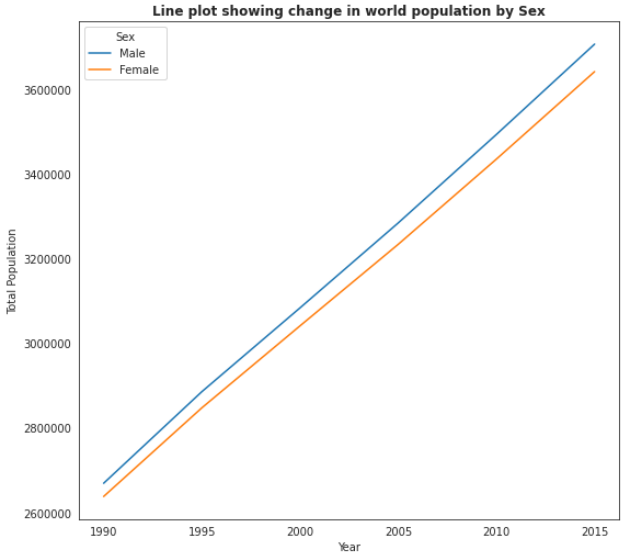


Figure 2.2 shows that there are more males than females in the total world population over the period of time under study. This suggests that the sex imbalance in international migration might not exist. The following sections will further examine the dataset in order to come to a conclusion.

Figure 2.2 Total population by Sex, 1990-2015



Is sex imbalance a worldwide trend or is it just a global level trend? To look at this topic from a different level, I sliced the dataset by continent. Figure 2.3.a is a small multiple of line charts showing the migration trend by continent (Asia, Europe, Africa, Oceania, Latin America, and Northern America). Green represents female migrants; orange represents male migrants, and blue represents both sexes which shows the global trend.

The charts suggest that the sex imbalance in international migration varies by continent.
In Asia and Africa, there are more male migrants than female migrants, which aligns with the global level trend. However, in Europe, the pattern is reversed with more female than male migrants. In Oceania, Latin America, and Northern America, the number of male and female migrants is more balanced, with a slight increase in the number of female migrants in recent years.

Figure 2.3.a. Migration Trend by continents

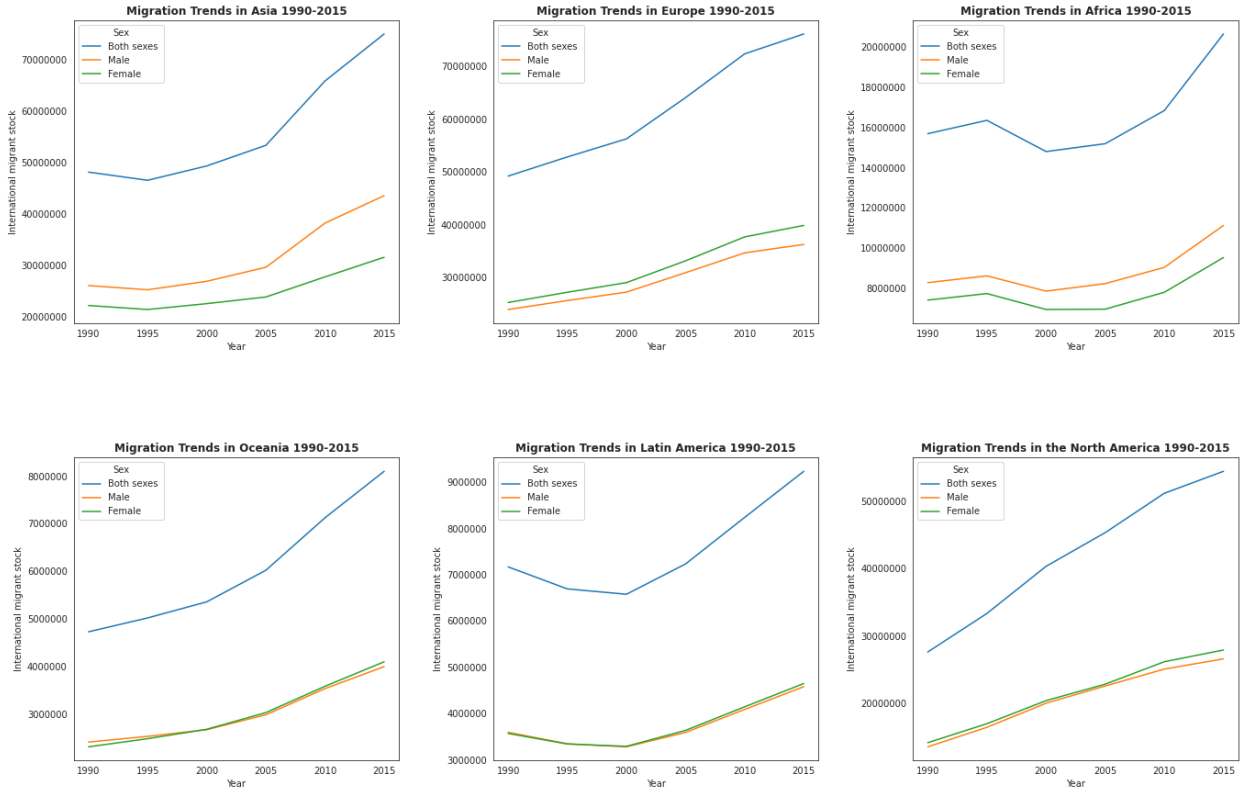


Figure 2.3.b. suggests a consistent trend of more female migrants moving to developed regions and more male migrants moving to developing regions. A hypothesis derived from this analysis is that more males migrate to developing countries for labour work opportunities, further research can be done to examine this hypothesis.

Figure 2.3.b. Migration Trend by Regions

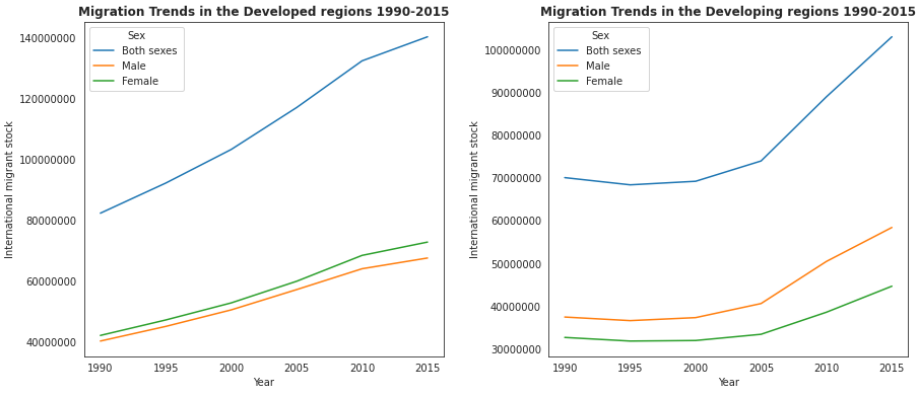


Figure 2.4 is a violin plot showing the international migrant stock as a percentage of the total population between developed and developing regions and aims to compare the differences between female and male migrants. The chart suggests a similar idea that there is a higher percentage of female migrants in

developed regions and a higher percentage of male migrants in developing regions. The plot shows that the percentage of migrants in developed regions is around 9%, which is higher than the percentage of migrants in developing regions, which is around 2%.

Figure 2.4 International Migrant Stock as a Percentage of the Total Population by Sex

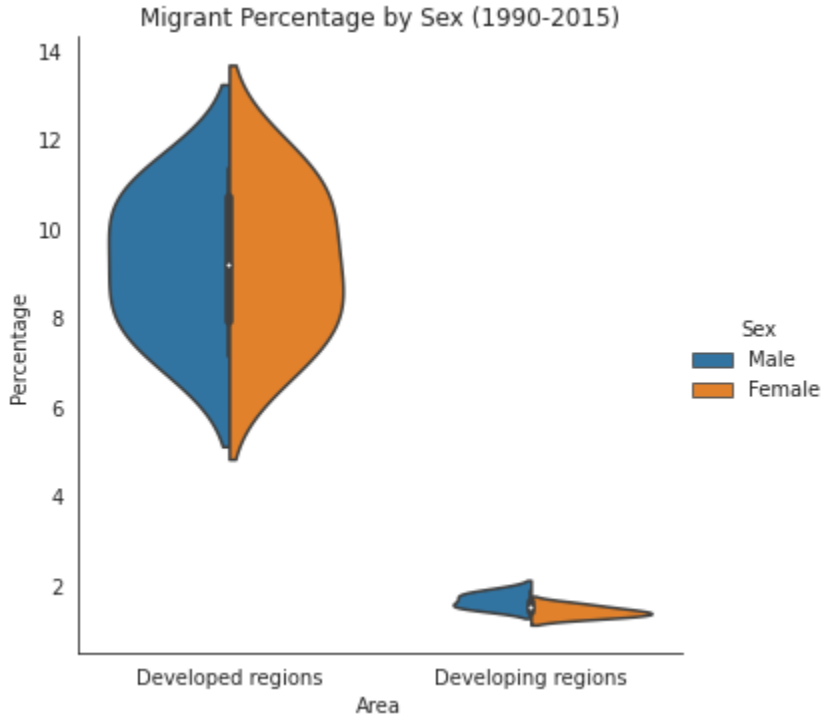


Figure 2.5 is a small multiple of line charts that compares female migrant percentages in each continent.

In Asia and Africa, there are fewer female migrants than male migrants, with the percentage continuing to decline over time. This could potentially be due to the fact that major Asian and African developing countries need labour workers, and males may be more suitable for these roles.

In contrast, Europe, Northern America, and Latin America have a female migrant percentage slightly over 50, indicating that there are more female migrants in these continents. The percentage of female migrants in Europe and Latin America consistently increases from 1990 to 2015, while in North America the percentage of female migrants declined from 1990, reached its lowest point in 2005, and then increased back to around the same level as 1990. In Oceania, the percentage of male and female migrants is about the same.

Figure 2.5 Female Migrant Percentage by Continents

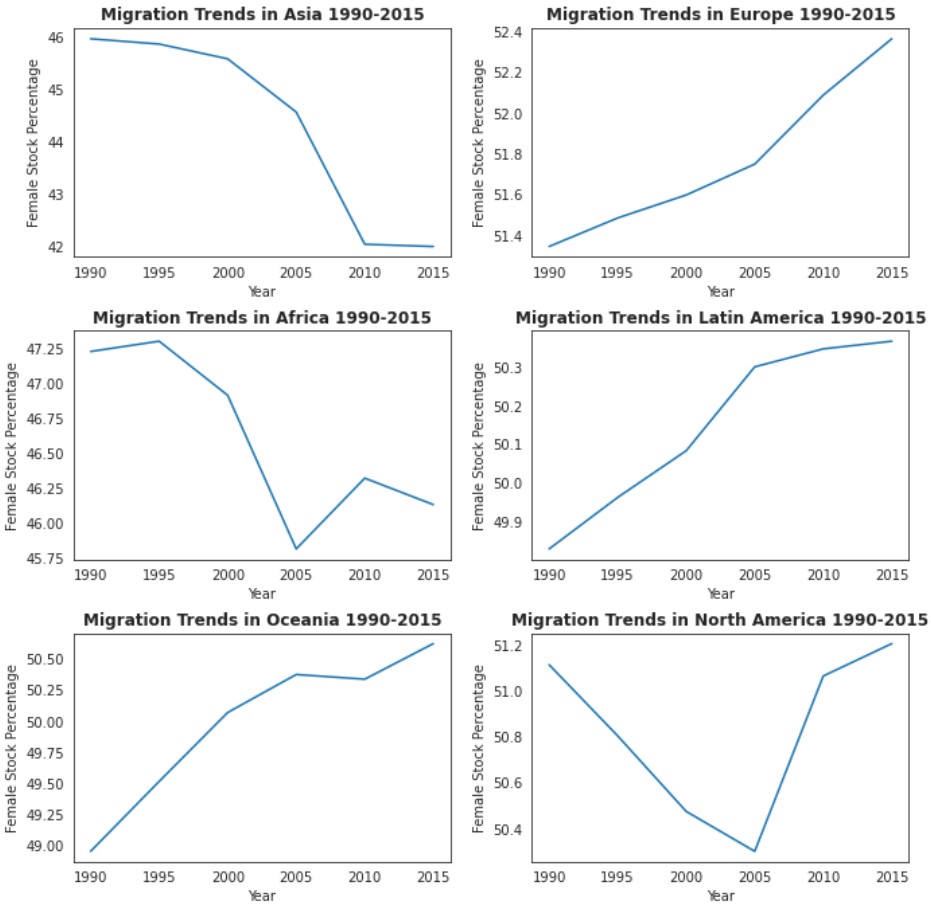
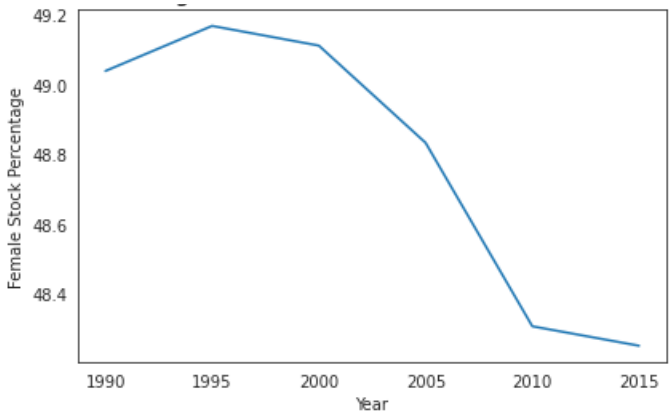


Figure 2.6 shows that overall, there are consistently fewer female migrants (a percentage less than 50) compared to male migrants on a global level. It also indicates that the percentage of female migrants increased in 1995 and almost balanced with male migrants, but then began to decline after that point.

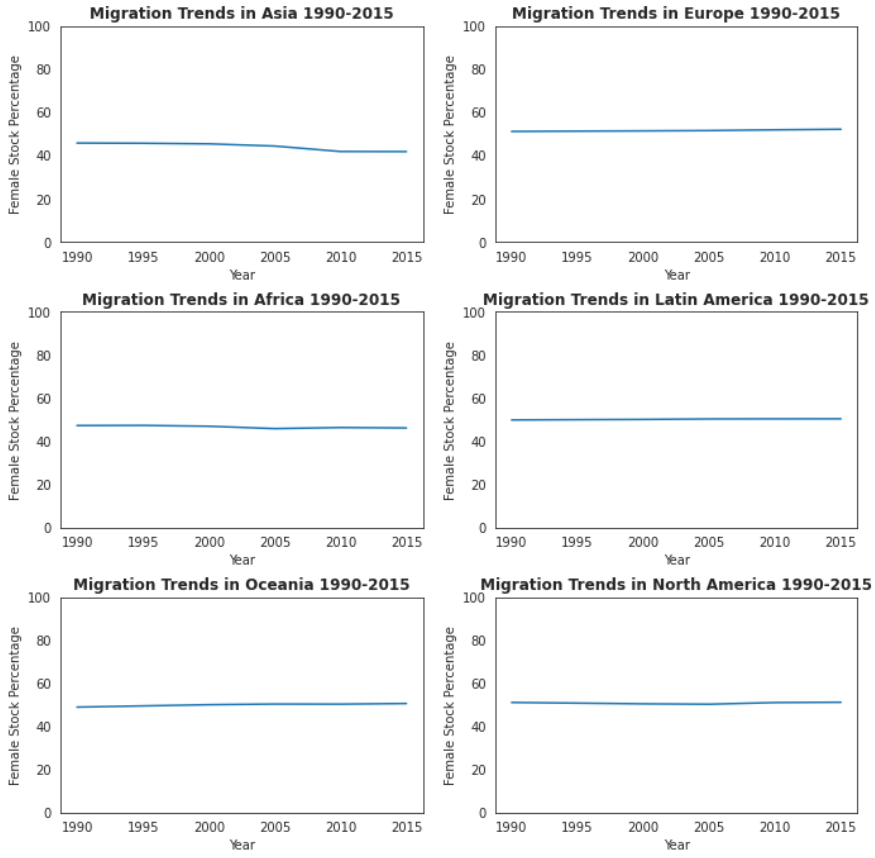
Figure 2.6 Female Migrant Stock Percentage in the World, 1990 - 2015



The line charts in figure 2.5 presented drastic changes in the percentage of females across the continents, but when the y-axes for the line charts were set to the same range of 0-100, representing the percentage, the change in the percentage of female migration across the continents becomes almost unnoticeable. In

Figure 2.7, the slope presenting the changes in the percentage of female migrants across the continents is relatively flat, with only Asia and Africa having a noticeable difference in the slope.

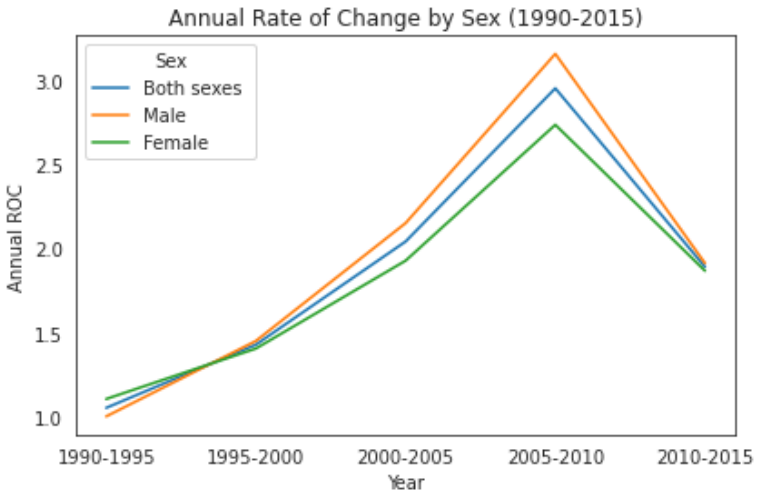
Figure 2.7: Female Migration Percentage by Continents (adjusted Y-axis)



Visualization has the power to tell different stories depending on how the data is presented. Figure 2.5 and Figure 2.7 present the exact same dataset with different y-axes, resulting in two different interpretations of the data. Figure 2.5 may lead one to conclude that there has been a drastic change in female migration percentages across continents, while Figure 2.7 suggests an almost unnoticeable change. According to Tufte's principle about graphical integrity, Figure 2.5 could be considered misleading because it may not accurately represent the data.

Figure 2.8 shows the annual rate of change in migrant stock across the globe, with positive numbers indicating an increase. Overall, the migrant stock increased, but there is a slower rate of growth from 2010 to 2015 compared to the previous time period. **This trend does not differ between the sexes.** Starting in 2010, while international migration continues to grow, the growth rate has decreased.

Figure 2.8 Annual Rate of Change in migrant stock by Sex



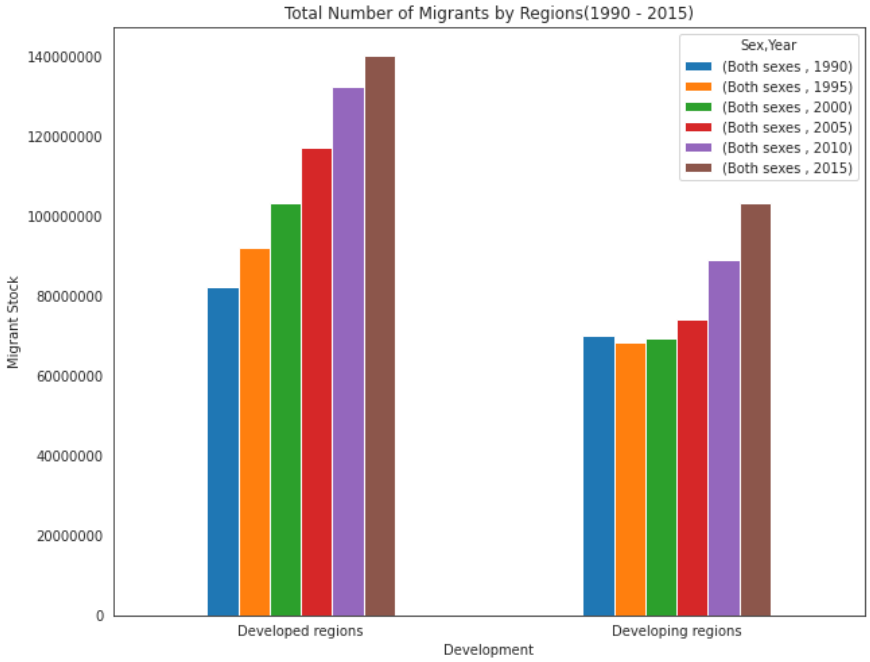
After analyzing the dataset using various graphical charts, the report concludes that there is a difference in the international migration trend between male and female migrants, with certain regions or continents attracting more males than females and vice versa. However, this does not necessarily mean that there was a sex imbalance in international migration over the time period from 1990 to 2015.

3. Overall Trends in International Migration

Research question: What are the overall trends in international migration?

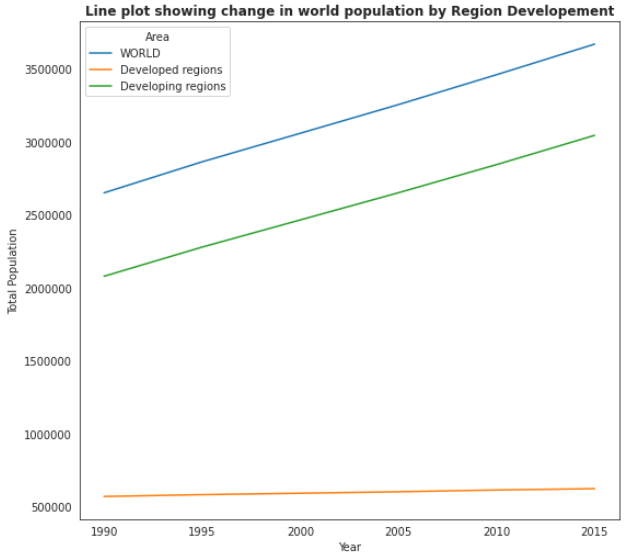
Figure 3.1 is a multiple bar chart comparing the number of international migrants between developed and developing regions from 1990 to 2015. It shows that the number of international migrants has increased over time for both developed and developing regions. However, in the developing regions, there is a slight drop in international migrants from around 1995 to 2000.

Figure 3.1 Numbers of International Migrants by Development, 1990-2015



An interesting finding is that in developed regions, while the number of international migrants increases over time, its total population remains almost the same from 1990 to 2015. As shown in Figure 3.2, the orange line presenting the developed regions is nearly flat across the years. A hypothesis is that more people over the years migrated from one developed country to another developed country, thus the total population barely changed while the number of migrants increased. If we have data on migration origins, we may be able to accept or reject this hypothesis.

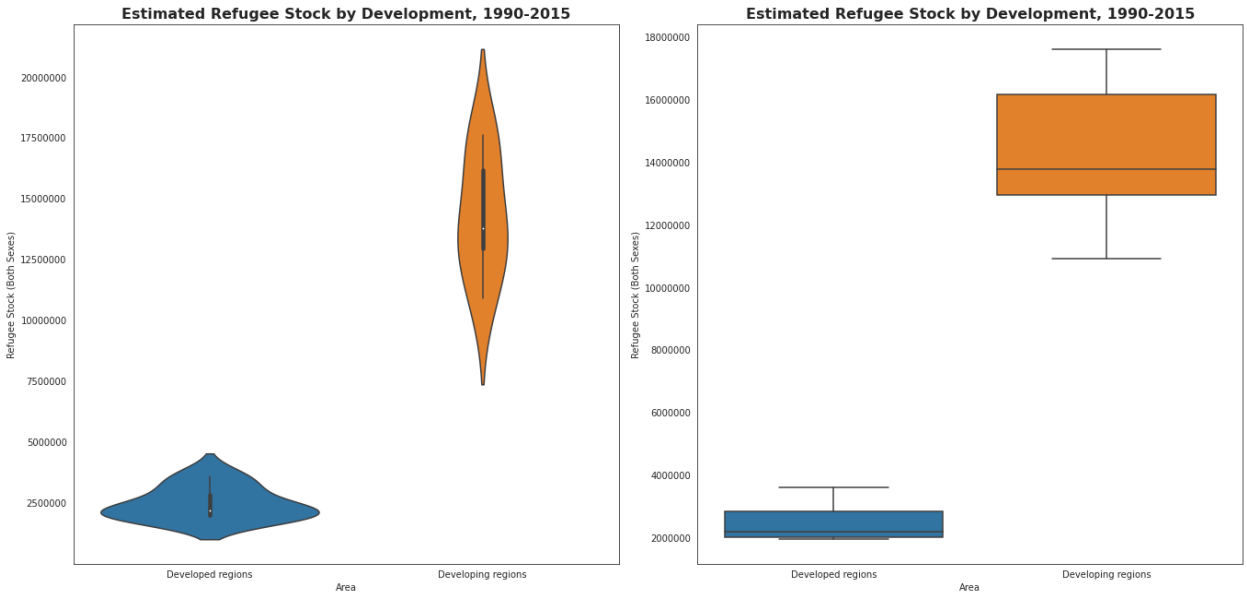
Figure 3.2 Change in World Population by Region, 1990-2015



3.1 Refugee

Refugee migration is a common reason for international migration. The violin plot and box plot in Figure 3.3 show the distribution of estimated refugee stock in both developed and developing regions over the time period from 1990 to 2015. There is a clear indication that there are more refugees moving to developing regions compared to developed regions.

Figure 3.3: Estimated Refugee stock in the world from 1990 to 2015



In the violin plot, Developed regions has a shape that is very wide in the middle, which indicates the refugee stock across the years is highly concentrated around the median. Whereas for Developing regions, the refugee stock of the selected period of time is more spread apart.

The box plot presents a similar visualization. Developed regions is positively skewed with the median closer to the lower quartile. It suggests that from 1990 to 2015, most years have refugee stocks around a similar range, with the exception of a high refugee stock year causing the upper whisker a lot longer than the lower whisker. For Developing regions, the lower and upper whiskers are around the same length, it is still slightly positively skewed with the median closer to the lower quartile. This indicates that from the minimum to Q1 and from the maximum to Q3 are around the same distance.

Figure 3.3 concludes that developed and developing regions have a different range of estimated refugee stock around the period of time under study. But to understand the trends across the years, line charts are created.

Figure 3.4 shows the estimated refugee stock by development from 1990 to 2015. Comparing the data from 1990 to the data from 2015, there is a slight increase in refugee stock for developing regions; and a slight decrease in refugee stock for developed regions.

Figure 3.4

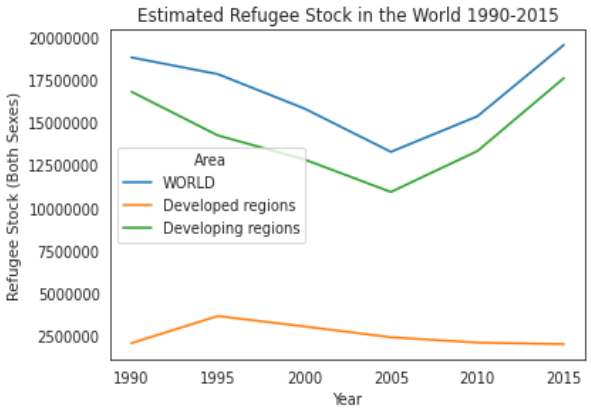


Figure 3.5

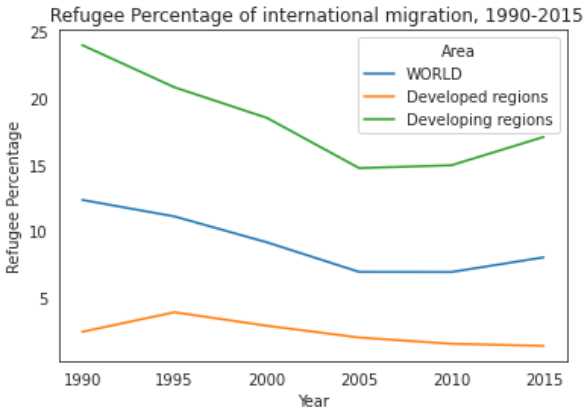


Figure 3.4 shows a decline in refugee stocks among international migrants in developing regions from 1990 to 2005, followed by an increase starting in 2005 and continuing until 2015. Figure 3.5 which shows the percentage of refugees among international migrants in developing regions from 1990 to 2015, does not consist of a sharp increase in the percentage of refugees in developing countries starting in 2005.

This suggests that **the increase in international migration in developing regions during this time period was not limited to refugees, but included other types of migrants as well.** To confirm, we can refer to Figure 2.3.b (right chart) which shows the migration trend in developing regions with a visible increase starting in 2005.

Figure 2.3.b. Migration Trend by Regions

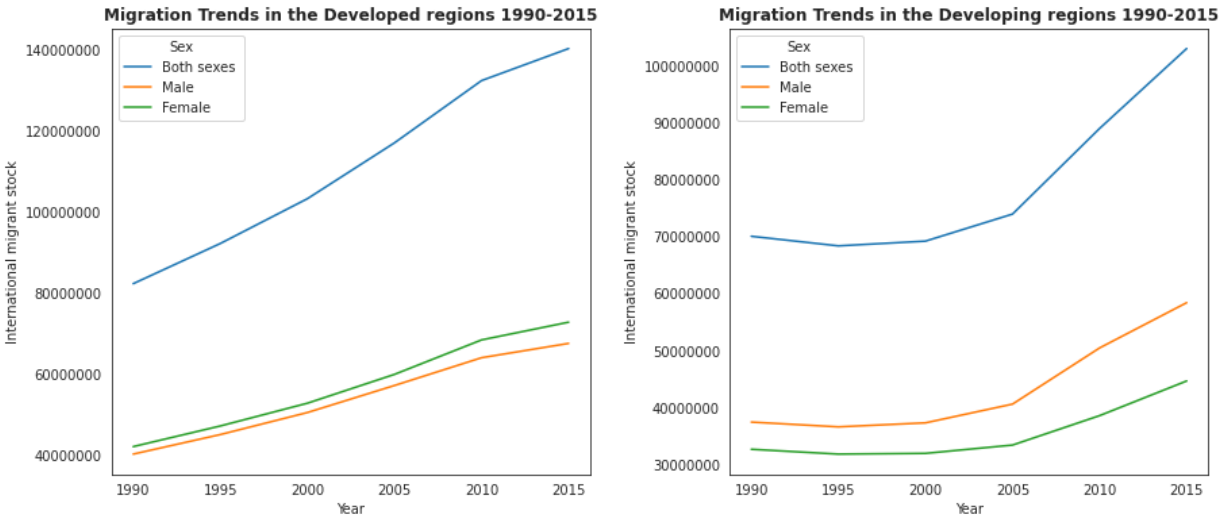


Figure 3.4 indicates that there was an increase in the number of refugees in developed regions in 1995, followed by a decline afterwards. This aligns with the positively skewed box plot in Figure 3.1 that the increase in refugee stock in 1995 is reflected by the long upper whisker. When we compare the number of refugee migrants in developed regions to overall international migration in these regions, Figure 2.3.b (left chart) shows a gradual increase in international migration in developed regions from 1990 to 2015

without any decline. This suggests that **while the overall number of migrants to developed regions has increased, the number of refugees among these migrants has decreased over time.**

4. Applying Visualization Principles

The visualizations presented in this report follow a number of Tufte's visualization principles. First, all charts presented in this report serve a clear purpose, which is to answer the research questions stated in the Overview section.

To avoid Chartjunk, or the usage of unnecessary graphical effects, this report applied bar charts, line charts, violin plots, and boxplots for graphical analysis because they can intuitively present the dataset.

This report values the importance of graphical integrity and some actions were done to follow the principle. For example, the charts presented all used labellings that are directly proportional to the numerical quantities represented. The y and x-axes are labelled with proper titles that can explain the data. Additionally, Section 2 discussed how the small multiples of line charts showing the female migration trend by continents could be misleading since each chart has a different y-axis value. To avoid this issue, I created a second version of the small multiples but applied the same y-axis value for all charts.