# A Visual Representation of Trends in International Migrant Stock（2015 Version)

by
Laien Zhou (1005490434)

A final paper submitted in conformity with the requirements
for the course of INF1340H

Faculty of Information
University of Toronto

# 1. Introduction

This paper focuses on the visual representation of Trends in International Migrant Stock, where the dataset is from the Department of Economic and Social Affairs of the United Nations. The data is collected in 2015 and mainly focuses on the period from 1990 to 2015. The presentation of this article follows Tufte's visualization principles including showing data by comparisons, demonstrating how variables impact others, combining various data, incorporating multiple modes of information, being attached with detailed titles and describing the state and trend. The visual representation of the variables in this paper will be consistent with the order in the original material

# 2. Visual representation and detailed analysis

## 2.1 International migrant stock at mid-year

The sheet 1 of the table is firstly read as a dataframe by pd.read_csv() and then meaningless lines on the top of the table are dropped and column names are reallocated to fit the corresponding data. Rows with "na" values are dropped. The types of values of each cell is reallocated as "numeric" to ease the further calculation and analysis. All rows that do not belong to a specific country are dropped. The data is then sorted by the values of 1990,and is then transformed to table form.

First ten rows of the data are picked alone and then figure 1 illustrates the international migrant stock at mid-year of the top 10 countries from 1990 to 2015, 5 years apart.
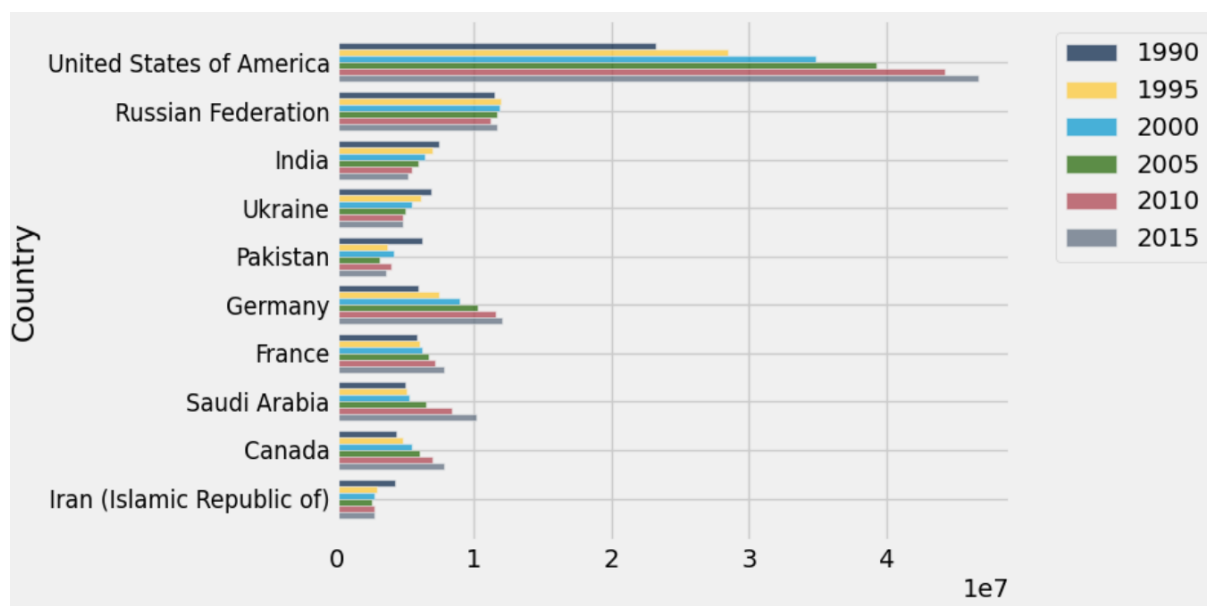


figure1: The International migrant stock at mid-year of the top 10 countries

From the figure above it is observed that the U.S. is far outnumbered, while the gap between countries other than the U.S. is at a relatively small level. Among the 6 years that data was collected, the U.S. always stayed in the first place in the international migrant stock at mid-year

Then the data is melted in gender and year and boxplots of both selected top 10 countries and all countries are created, which separates on gender.

From the left graph (figure 2) it is illustrated that the deviations of international migrant stock at mid-year of both male and female increase in the time period, showing that the number of immigrants between different countries was widening, and the immigration numbers are shifting to a few

countries, especially to the U.S. Figure 3 on the right also demonstrates the same thing: the international migrant stock at mid-year of most countries hovers around 0, and international migrant stock at mid-year of the U.S. stays at a high level for both male and female and keeps going up in the time period.
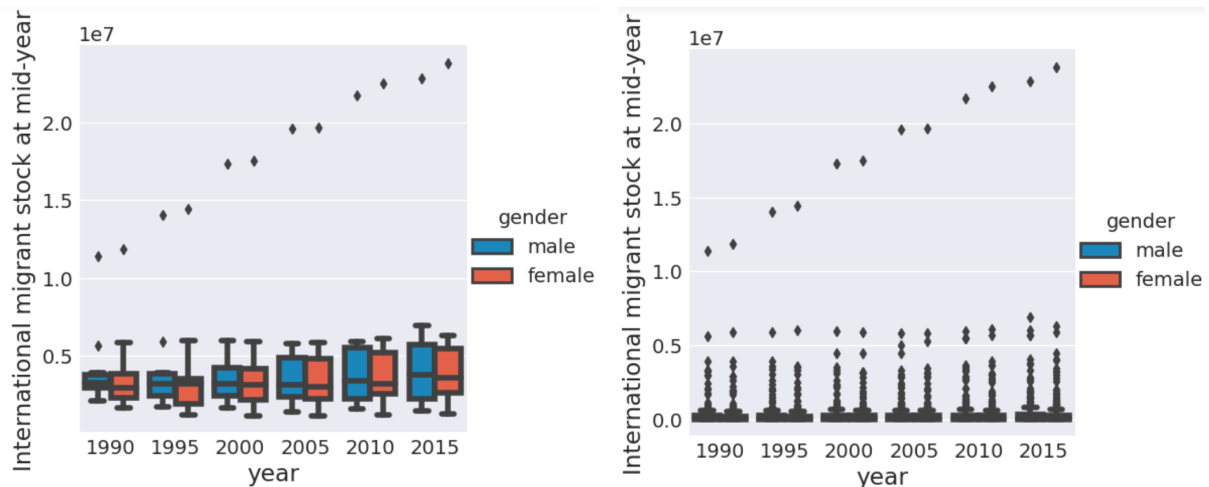


figure 2 (left): Boxplot of International migrant stock at mid-year of the top 10 countries by gender
figure 3 (right): Boxplot of International migrant stock at mid-year of all countries by gender

## 2.2 Total population of both sexes at mid-year

The sheet 2 of the table is read as a dataframe by pd.read_csv() and then meaningless lines on the top of the table are dropped and column names are reallocated to fit the corresponding data. Rows with "na" values are dropped. The types of values of each cell is reallocated as "numeric" to ease the further calculation and analysis. All rows that do not belong to a specific country are dropped. Then the data is sorted by the value of 1990,and is then transformed to table form.

First ten rows of the data are picked alone and the corresponding graph is created as figure 4, which
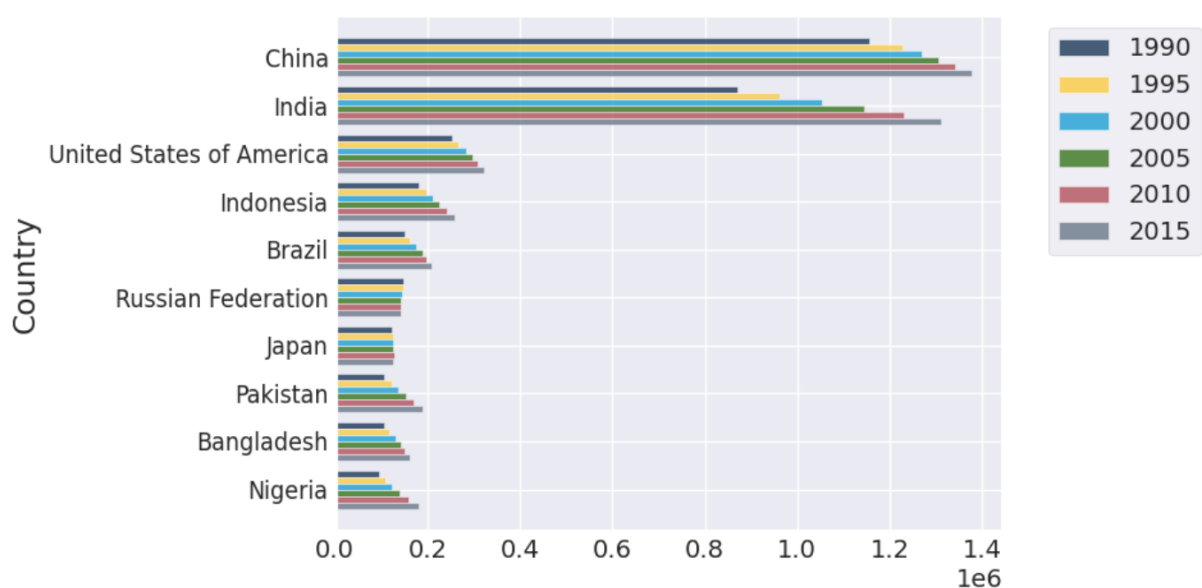


figure 4: Total population of both sexes at mid-year of the top 10 countries

shows the distribution of the top 10 countries in the total population of all genders at mid-year, the populations of China and India are far greater than other countries', which is followed by the U.S. that has the higher international migrant stock.
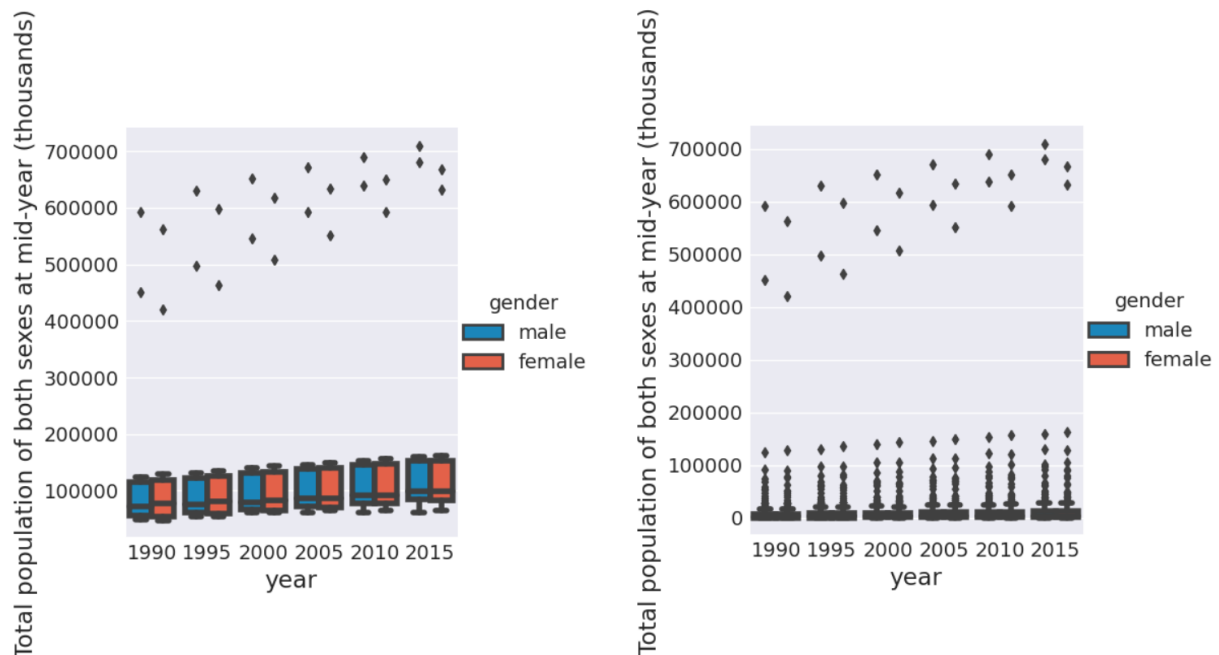



figure 5: Total population of both sexes at mid-year of the top 10 countries by gender
figure 6: Total population of both sexes at mid-year of all countries by gender

The graph above illustrates the boxplot visualization of the total population of different genders at mid-year in selected top 10 countries (on the left) and in all countries (on the right). We see that the percentage of male and female population in most countries does not change while the populations of both genders keep growing at a smooth rate.

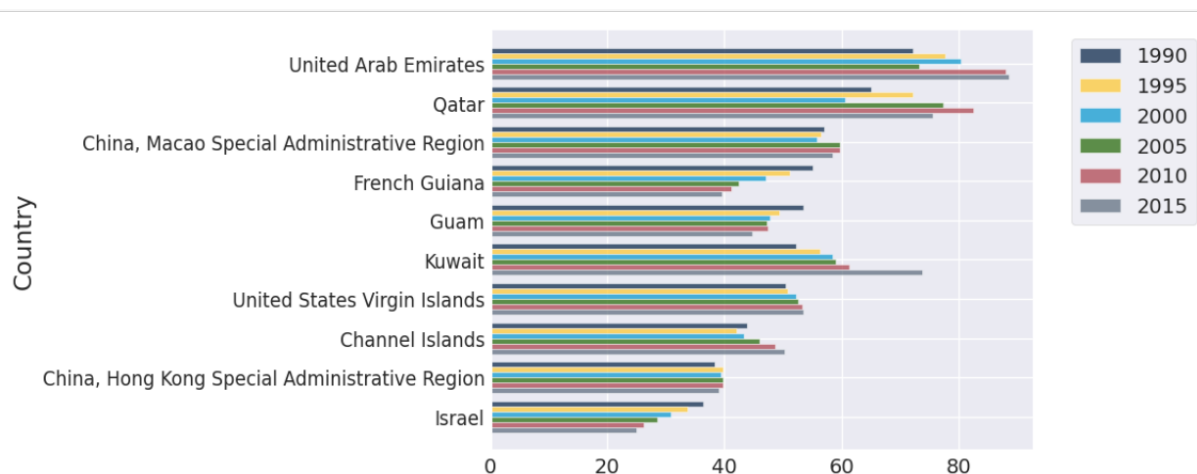## 2.3 International migrant stock as a percentage of the total population



figure 7: International migrant stock as a percentage of the total population of top 10 countries

The sheet 3 of the table is read as a dataframe by pd.read_csv() and then meaningless lines on the top of the table are dropped and column names are reallocated to fit the corresponding data. Rows with

"na" values are dropped. The types of values of each cell is reallocated as "numeric" to ease the further calculation and analysis. All rows that do not belong to a specific country are dropped. Then the data is sorted by the value of 1990, and is then transformed to table form. First ten rows of the data are picked alone and the corresponding graph is created as figure 7.

From figure 7, it is illustrated that the percentage of international migrant stock in the total population of the top 10 countries does not differ a lot from each other, and most of them do not appear in the top 10 countries of international migrant stock in figure 1.
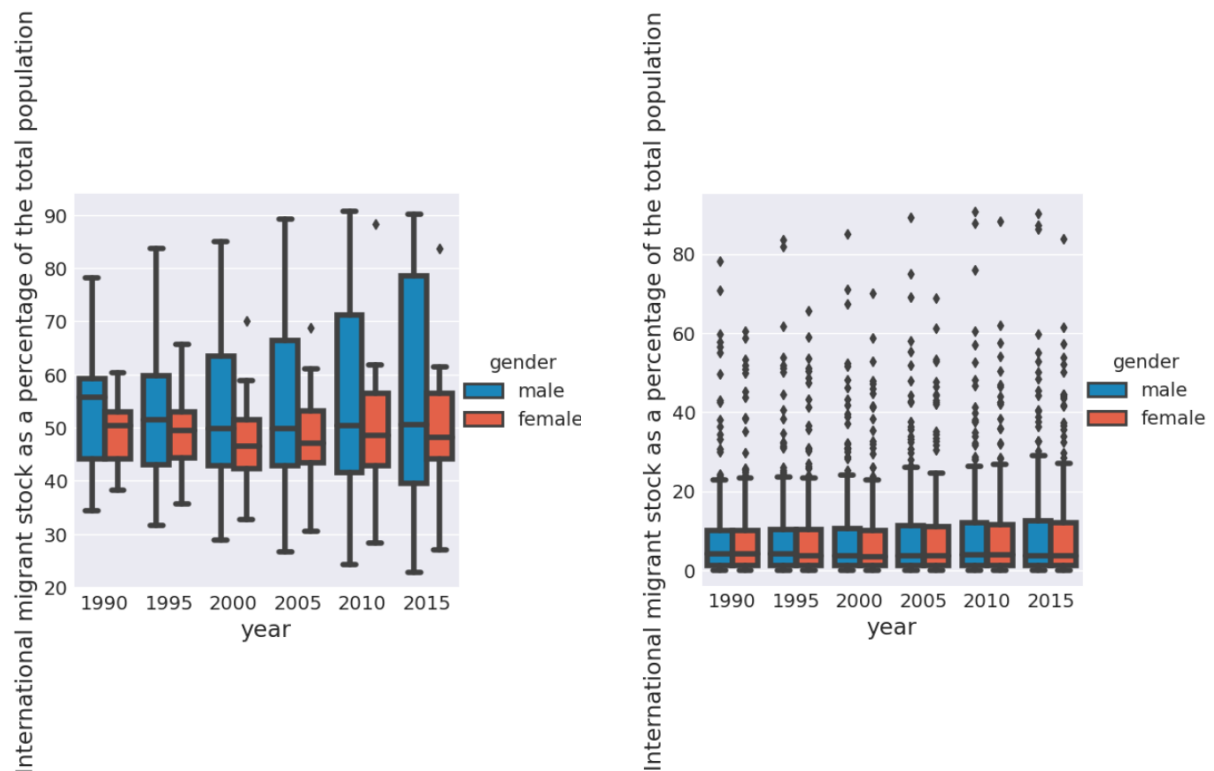


figure 8: International migrant stock as a percentage of the total population by gender in top 10 countries
figure 9: International migrant stock as a percentage of the total population by gender in all countries

Figure on the left (figure 8) shows the boxplot visualization by gender of the top 10 countries in the international migrant stock as a percentage of the total population. It is significantly shown that among the top 10 countries the distribution of male's percentage of international migrant stock in the total population changes a lot as the deviation keeps going up in the time period, and the female's percentage of international migrant stock in the total population stays in a smooth and relatively concentrated distribution. The figure on the right (figure 9) shows that the distribution of the percentage of international migrant stock in the total population of most countries in the world stays at a low level, illustrating that the phenomenon of the international migrant stocks is influenced by a few top countries.

## 2.4 Percentage of the international migrant stock by gender

The sheet 4 of the table is read as a dataframe by pd.read_csv() and then meaningless lines on the top of the table are dropped and column names are reallocated to fit the corresponding data. Rows with "na" values are dropped. The types of values of each cell is reallocated as "numeric" to ease the further calculation and analysis. All rows that do not belong to a specific country are dropped.

The data is sorted by the value of 1990 (female), and then the data of male is appended as the complementary set of female.

Then the column names of the dataset is reallocated and boxplot is created by gender as figure 10.
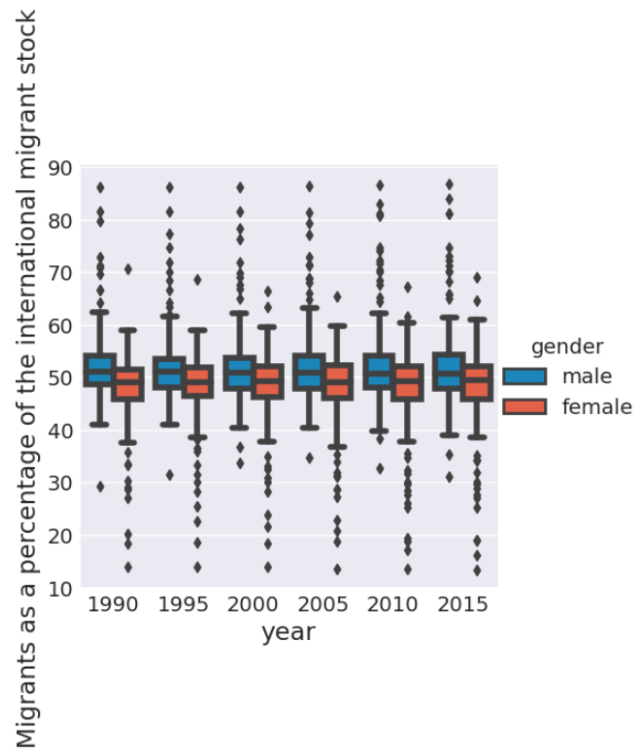
The figure on the left illustrates that in most countries in the world, the percentages of male's and female's international migrant stock keeps in a similar position, the mean of male's percentage is a few higher than the mean of female's, and the deviations of both male's and female's percentage keep approximately the same.

figure 10: Percentage of the international migrant stock by gender

## 2.5 Annual rate of change of the migrant stock

With the similar method, figure 11 shows the annual rate of change of the migrant stock of the top 10 countries, where the top countries are selected based on the annual rate of change of the migrant stock between 1990-1995, which is the beginning period of the data.
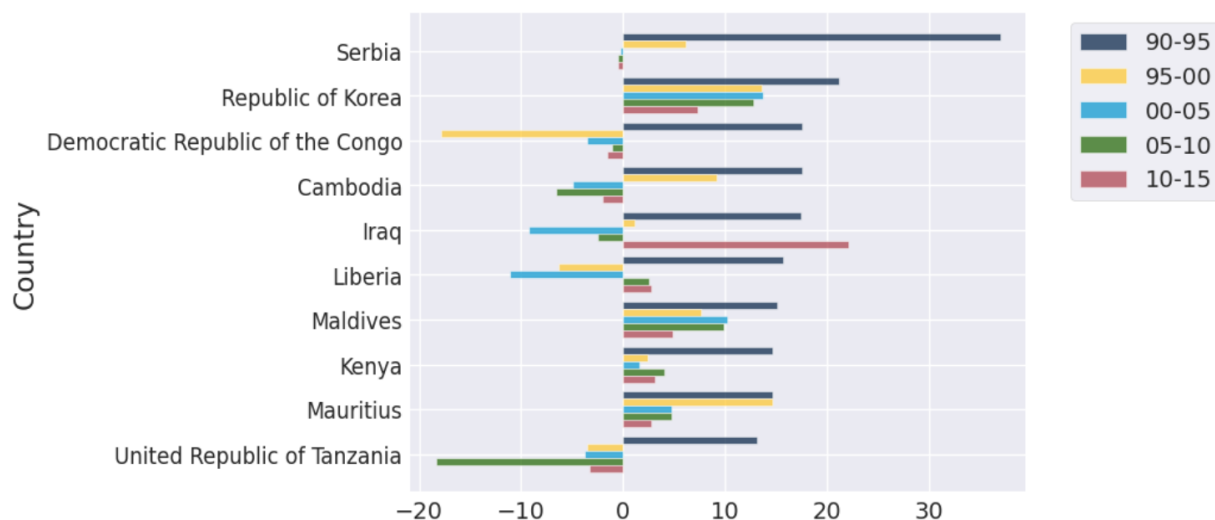
figure 11: The annual rate of change of the migrant stock of the top 10 countries (90-95)

It is observed that among most of the countries above, as their annual rate of change of the migrant stock in the first period is high, the annual rate in the next periods keeps going down, which reveals an attitude of those countries to prevent annual rate from keeping going up, whatever the phenomenon is caused by government or natural market regulation.
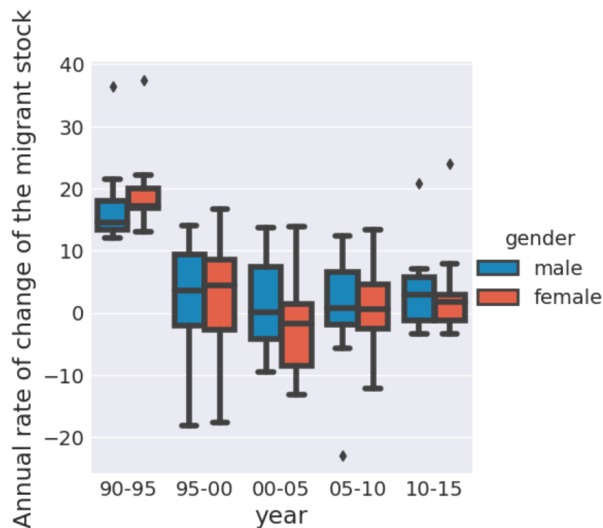


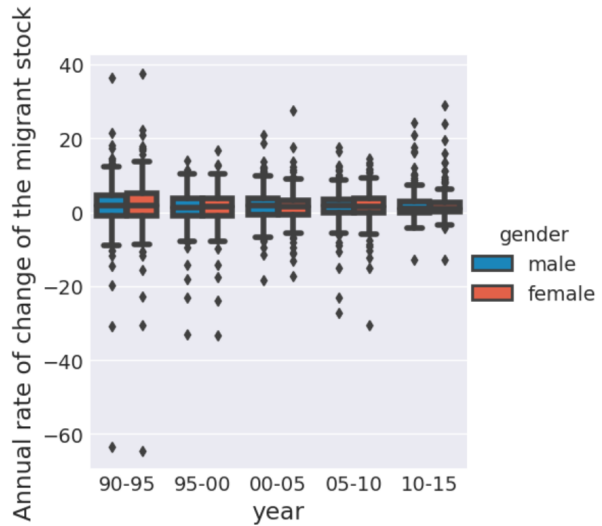figure 12: Boxplot of the annual rate of change of the migrant stock of the top 10 countries (90-95)
figure 13: Boxplot of the annual rate of change of the migrant stock of all countries

Figures above illustrate that the annual rate among the selected countries follows the phenomenon that a higher point in the first period leads to a continuous lowering in the subsequent period, while among all countries in the world, the trend of annual rate seems to be unchangeable in both the mean and the deviation.
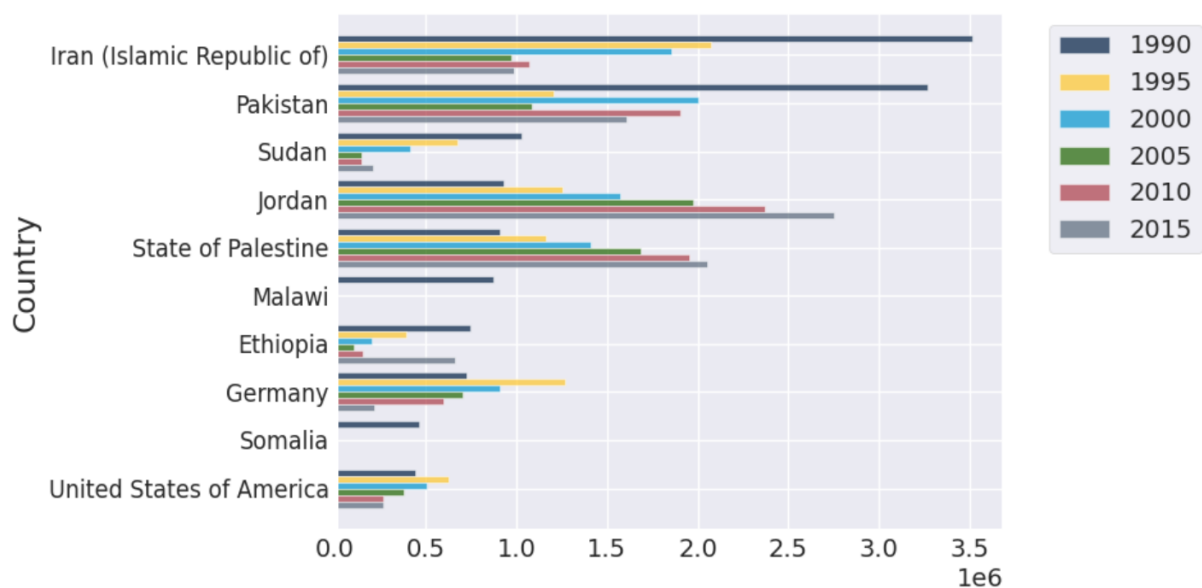
## 2.6 Refugee stock at mid-year



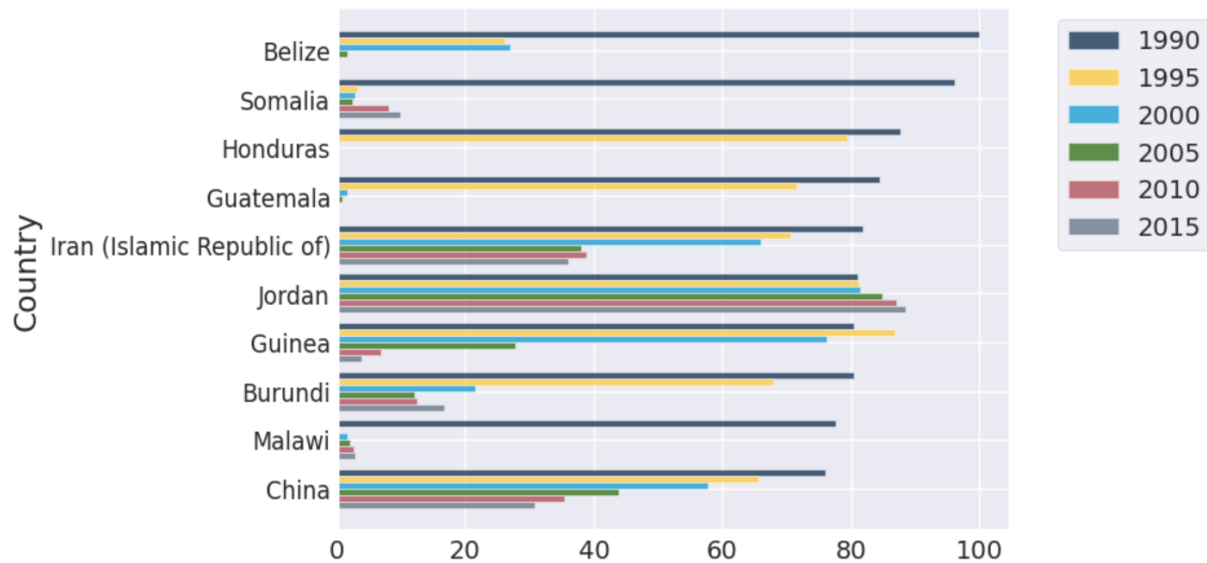figure 14: Estimated refugee stock at mid-year in top 10 countries

figure 15: Refugees as a percentage of the international migrant stock in top 10 countries
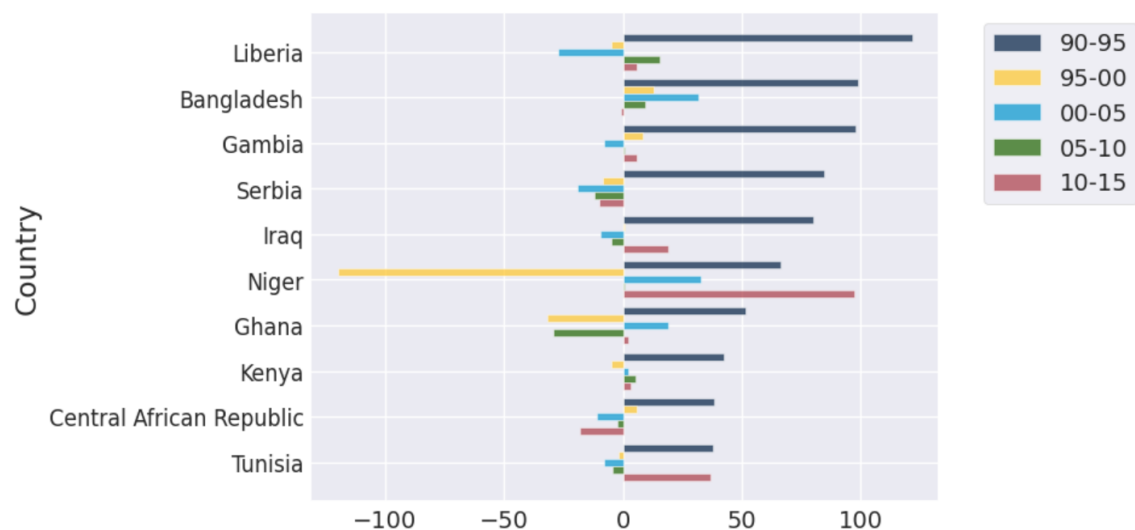


figure 16: Annual rate of change of the refugee stock in top 10 countries

It is illustrated that estimated refugee stock at mid-year goes down in countries that own a high amount of refugee stock in the first point (1990), as well as the percentage of refugees in the international migrant stock goes down in most of those countries, and the annual rate of change of the refugee stock also decreases rapidly in those countries who own a high annual rate of change of refugee stock in the first period.

# 3. Conclusion

The visualization of the data shows that for most countries, the amount of international migrant stock remains at a low level close to zero as well as the percentage of international migrant stock in the total population keeps at a low level. The annual rate of change of the migrant stock in most countries does not change and remains at a low level approximately at zero. The amount of refugee stock at mid-year of almost all countries is expected to decrease if the refugee stock of the country previously was at a high level, and the percentage of refugee stock and the annual rate of change of the refugee stock have the same pattern: the amount of the variable trends to go down rapidly once it stays in a high level at some point (in some period).

It is obviously shown that a few countries dominate the external appearance of the data of the international migrant stock, since that some countries often stay among the top 10 countries for various variables of international migrant stock. For instance, the U.S. stays in the highest position in the international migrant stock at mid-year among all countries, as well as its percentage of international migrant stock in total population and the annual rate of change of international migrant stock in most periods. The results are consistent with the fact that the U.S. is the largest immigrant country in the world and reflect the high international migrant stock attraction of the U.S. However, the U.S. seldom appears in the top 10 countries when we discuss with the figures of refugee stock, in other words, the U.S. does not host a lot of refugee stock while it keeps an attraction to international migrant stock. This might illustrate that the U.S. immigration management policy is more advanced than other countries, or it might be due to something that are not revealed in the data, and the exploration and study of related questions can be the direction of future attention.