**INF-1340 Final project**

**Cunrong Liu**

**Introduction：**

The main purpose of this project is to visualize tidy datasets (UN_Migrant Stock Total_2015) according to main principles. Explore the main trends of international migration through data visualization, by exploring the performance of different genders, different continents, and different years. Finally, discuss these phenomena and make a conclusion.

**Methods & Result：**

Table 1: international migrant stock at mid-year by sex and by major area, region country or area, 1990-2015.

The first table measures international migrant stock at mid-year by sex and by major area, region, country or area, during 1990-2015. I want to see the general trend of international migration, first I build a new data frame, filter "major area, region country or area" to only "world", "gender" and "year" remain unchanged. I want to see the general trend of international migration, first I build a new data frame, filter "major area, region country or area" to only "world", "gender" and "year" remain unchanged. According to tufte's principles, such content focuses on changes in world migration. Using this data frame, I can create a visualization. I choose bar plot because it can clearly compare the performance of different genders in this year. It can also effectively display two variables (year and gender) and clearly compare the number of immigrants in different years.

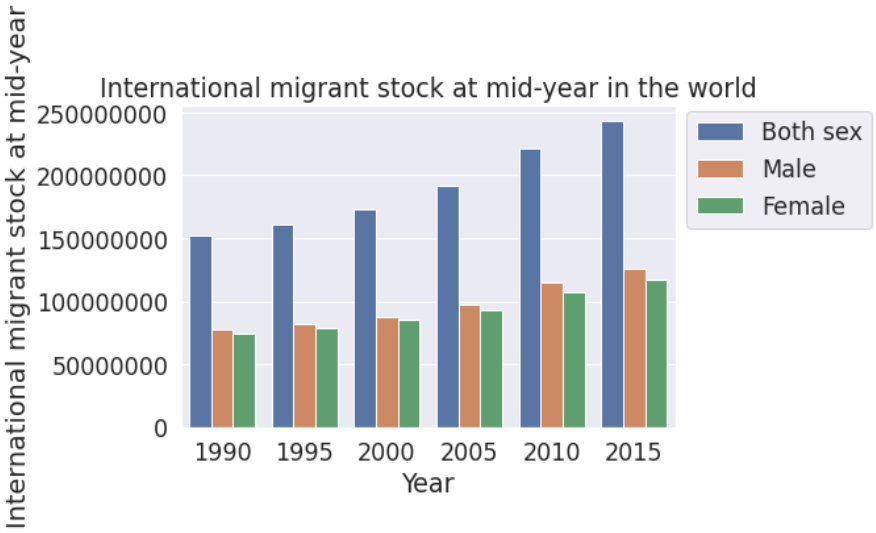


*Figure1: international migrant stock at mid-year in the world*

Through this graph, it can be seen that the number of immigrants has been rising every year, from 15 million in 1990 to nearly 25 million in 2015, and the gap between the number of immigrants between men and women is getting wider and wider.

Table 2: Total population at mid-year by sex and by major area, region, country or area, in 1990-2015(thousands)

The same method as the first figure, also create a new data frame, and then use the box plot to display the total global population, from 1990 to 2015, so that you can clearly compare the differences in the annual population, etc.

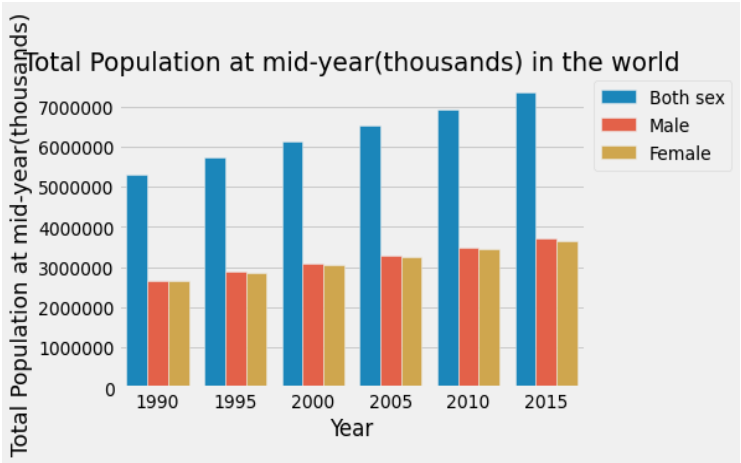


*Figure 2: Total population at mid-year (thousands) in the world*

In this figure, it can be seen that from 1995 to 2015, the total global population has been growing, not only "both sex", but also the total population of both male and

female. And the population of male and female is basically the same, there is no obvious difference, during 1990-2015.

The difference from before is that this time I created a data frame about the major continents, which include Asia, Europe, Africa, Oceania and Latin America. I want to explore the total population change for each continent between 1990 and 2015. I chose a line chart because it shows the trend of the population over time. Using small multiple time series can better reflect the changes in the relationship between variables as the indicator variables change

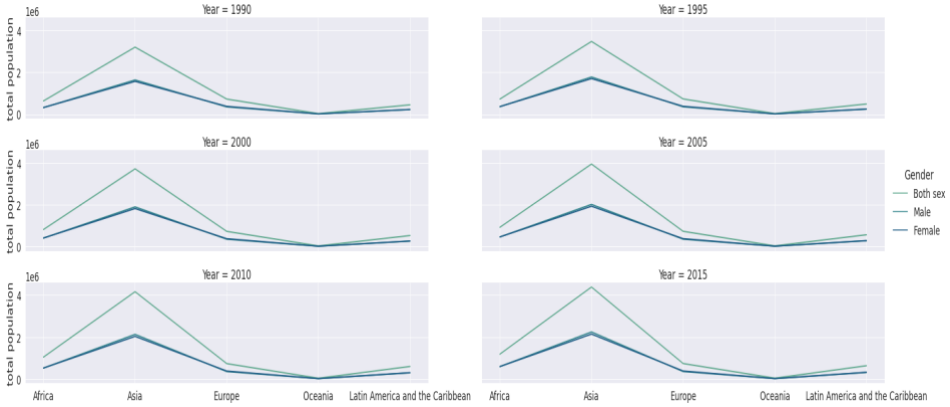


*Figure 3: e*

It can be seen that Asia, both male and female, had the largest total population between 1990 and 2015. Followed by Africa, Oceania has the smallest total population. And between 1990 and 2015, the total population of men and women in each continent was basically the same.

Table 3： International migrant stock as a percentage of the total population by sex and by major area, region, country or area, 1990-2015

Using the same method as the first figure, but this time use a line plot. I'll create a data frame looking at international migrant stock as a percentage of the total population by sex.

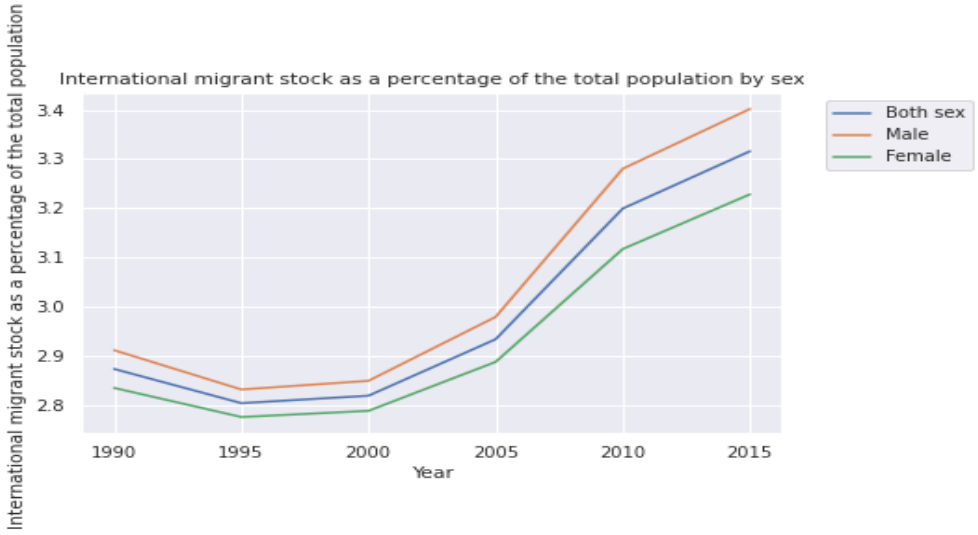


*Figure 4: International migrant stock as a percentage of the total population by sex*

It can be seen that for both male and female, the percentage of international migrants in the total population decreased between 1990 and 1995, but increased significantly in the following 20 years, especially after entering 2005. Male have always outnumbered female.

Table 4: Female migrants as a percentage of the international migrant stock by major area, region, country or area, 1990-2015

This figure is similar to the method of figure3, showing the situation of the major continent. In this graph, I use a line plot to compare the percentage of female immigrants in different continents as a percentage of international immigrants.

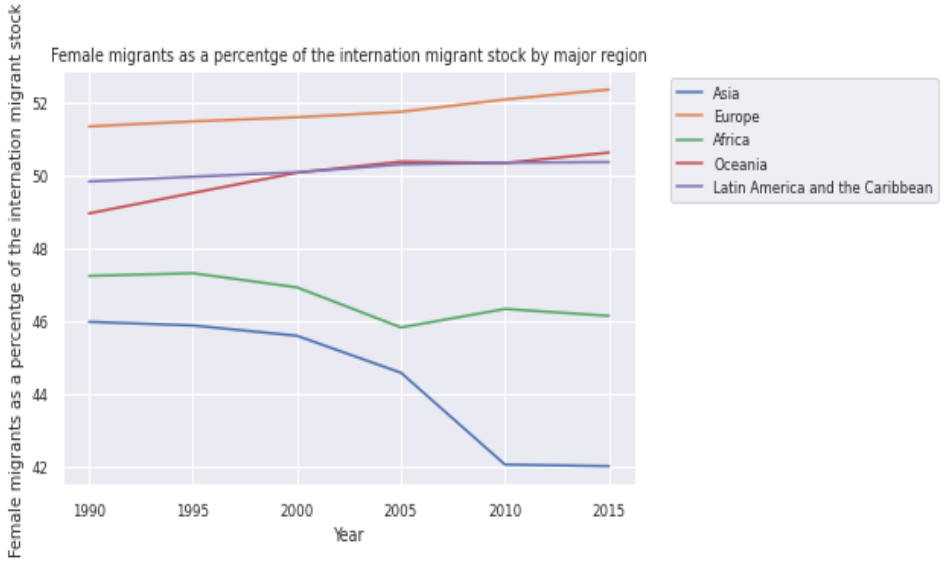

*Figure 5: Female migrants as a percentage of the international migrant stock by major region, 1990-2015*

It can be seen that the largest proportion of female immigrants is in Europe, and the least in Asia. Among them, women immigrants from Europe and Latin America account for more than 50% of the total immigrants in the continent.

Table 5: Annual rate of change of the migrant stock by sex and by major area, region, country or area, 1990-2015 (percentage)

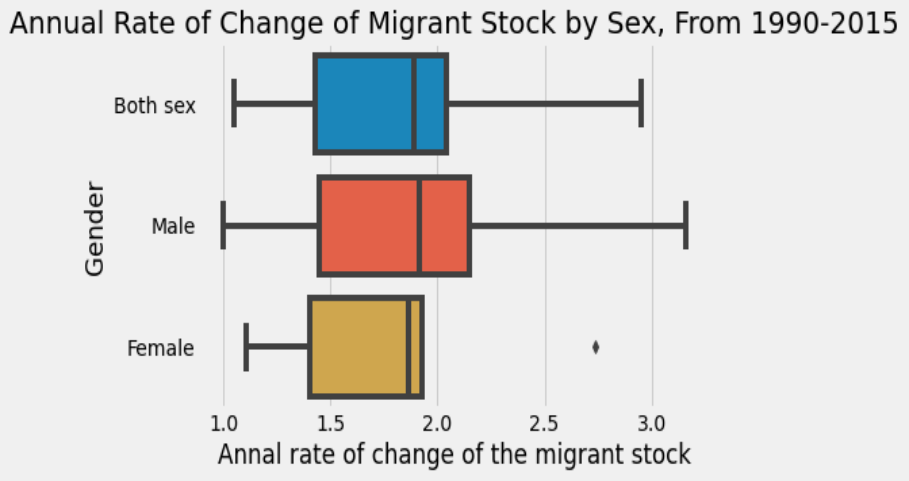


*Figure 6: Annual rate of change of the migrant stock by sex, from 1990-2015*

It can be seen that neither "both sex", "male" or "female" are normally distributed. Among them, the data distribution of "both sex" and "male" is wider. Female data has outliers

Table 6: Estimated refugee stock at mid-year by major area, region, country or area, 1990-2015

Same as the previous method, but I divided table6 into three parts:

1. Estimated refugee stock at mid-year
2. Refugees as a percentage of the international migrant stock
3. Annual rate of change of the refugee stock

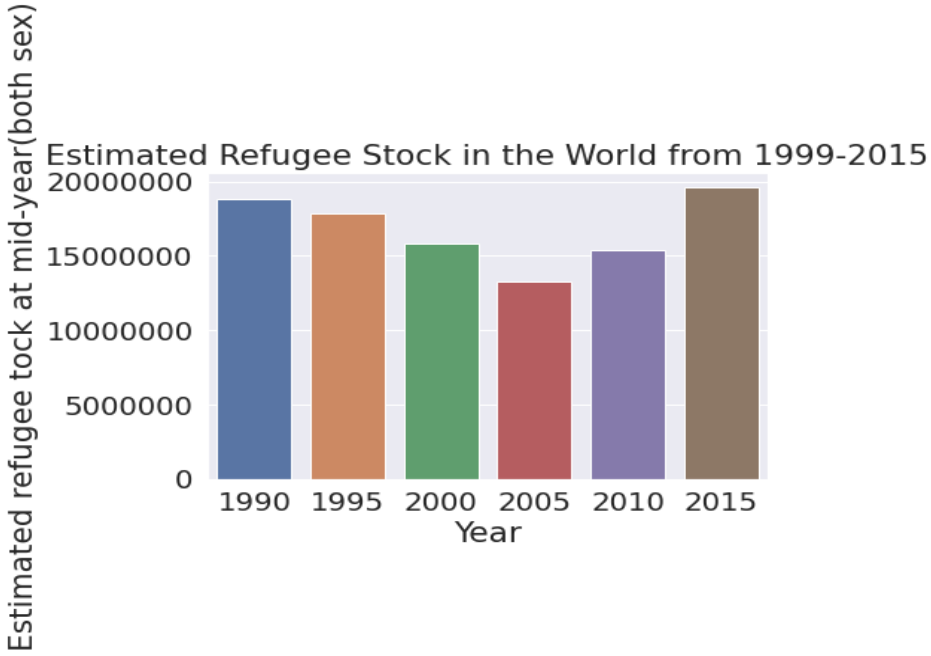


*Figure 7: Estimated refugee stock in the world from 1990-2015*

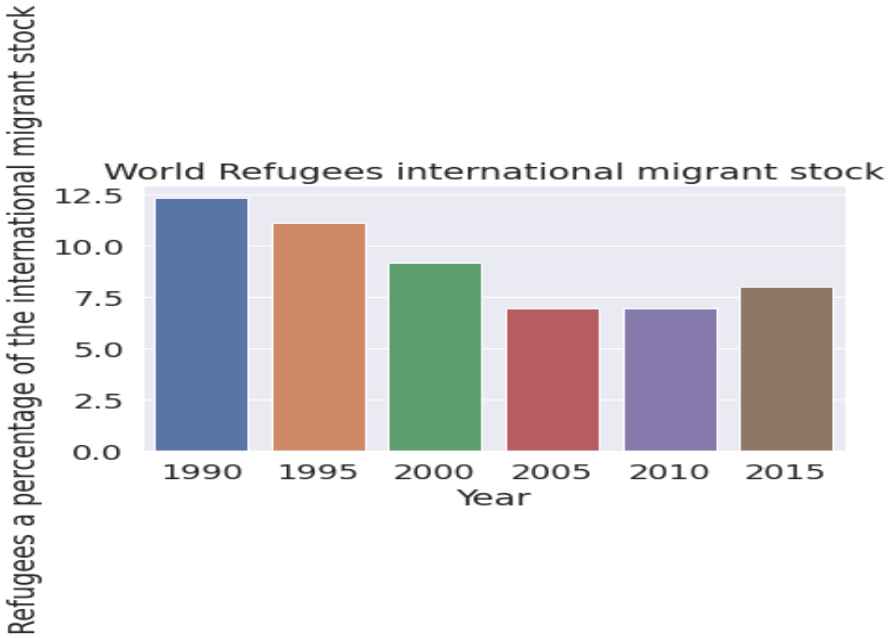


*Figure 8: world refugees international migrant stock from 1990-2015*

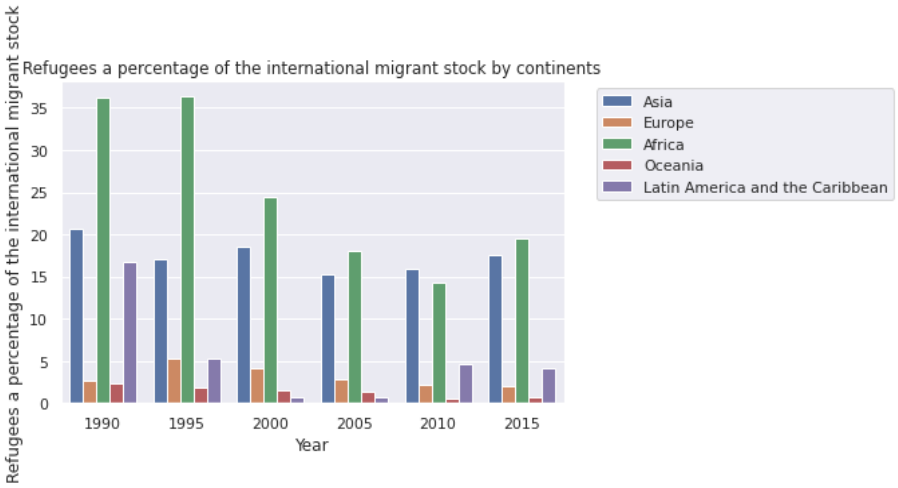


*Figure 9: Refugees a percentage of the international migrant stock by continents from 1990-2015*

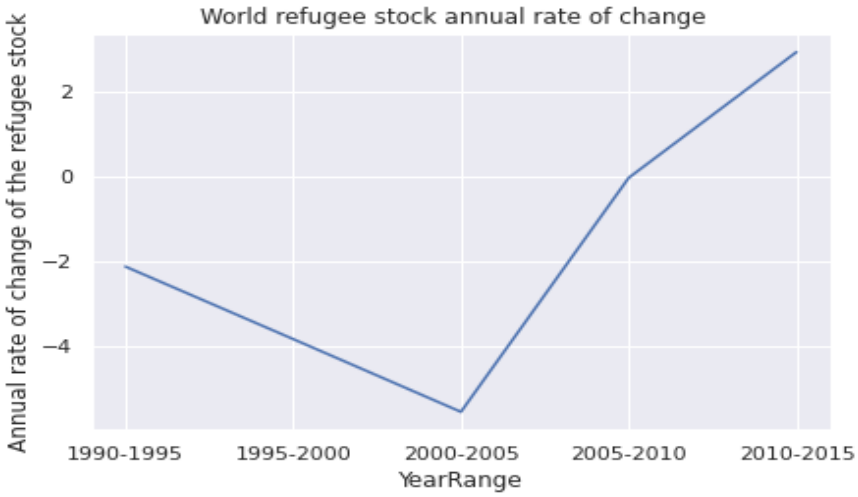


*Figure 10: world refugee stock annual rate of change from 1990-2015*

As shown in figure 7, from 1990 to 2005, the estimated refugee stock in the world decreased every year, reached the lowest point in 2005, then began to rise again, and the number of refugees is expected to peak in 2015. Similarly, from 1990 to 2005, world refugees international migrant stock began to fall, and then began to rise again. As can be seen from figure 9, refugees from Africa account for the highest percentage of international migrants on the continent, followed by Asia. Finally, from 1990 to 2005, the world refugee stock annual rate of change decreased year by year, but from 2005 to 2015, the world refugee stock annual rate of change increased sharply.

Discussion & Conclusion:

According to the above results, the number of immigrants is relatively small in

continents with a large proportion of developed countries, such as Europe, and the number of immigrants is large in continents with a large proportion of developing countries, such as Africa and Asia. This may be due to the fact that people tend to live more in developed countries, with higher economic levels and more immigration over time. Secondly, for refugees among immigrants, after entering 2005, there were some major international wars, such as the Iraq War and the Yugoslav conflict, which led to a large influx of refugees from Africa and Asia to Europe and the Americas.

In conclusion, Data visualization transforms large amounts of complex numerical data into more tractable visual representations. In this project, through data visualization, it is clear that from 1990 to 2015, the stock of international immigrants showed an upward trend, although the upward trend varies from region to region. Among them, whether male or female, Asia and Africa account for a particularly large proportion of immigrants.