



## School of Social Sciences and Philosophy Assignment Submission Form

Student Name:	Chenxi Li
Student ID Number:	23330541
Programme Title:	Applied Social Data Science
Module Title:	Spatial Data Analyst
Assessment Title:	<i>Assessment: Essay</i>
Lecture (s):	Prof. Martina Kirchberger
Date Submitted:	March 1, 2024

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: <http://www.tcd.ie/calendar>

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at <http://tcd-ie.libguides.com/plagiarism/ready-steady-write>

Signature: 

Date: March 1, 2024

## Spatial Data Analysis, Hilary Term 2024

### *Course Project (Max 1000 words)*

**Word Count: 785 words & 2 exhibits**

**Chenxi Li, 23330541**

My data set is migration data from the *immigrants, emigrants and Net Migration* section of the Migration Policy Instituion (MPI)<sup>1</sup>. Data from MPI is usually used for immigration research like immigrants trend (Stalker 2002, Helbling & Kalkum 2018), the relationship between poverty reductions and immigrants amount (Newland & Patrick 2004), etc. This data set is combined with 3 different spatial data, (1) the first is the net number of migrants by country from 1950 to 2020, (2) the second is immigrant and emigrant populations by country of origin and destination at mid-2020 estimates, (3) and the third is the total immigrant and emigrant populations by Country. The spatial data set (1) is raster data, (2) and (3) is vector data. The following table is the data overview and composition ideas used in this project.

In these three data, I found some potential vulnerabilities. I will divide my views into two parts, the discussion about the first data set – that is, raster data; and the second and third data set — that is, vector data. The following table is the data overview and composition ideas used in this project.

Num	Data Set Name	Type	Discuss
1	Net Number of Migrants by Country, 1950-2020	Raster	Part 1.
2	Immigrant and Emigrant Populations by Country, mid-2020	Vector	Part 2.
3	Total Immigrant and Emigrant Populations by Country	Vector	Part 2.

---

<sup>1</sup> Migration Policy Instituion, MPI, is an American liberal think tank which provides information of migration and refugee trends. The website of this data set is: <https://www.migrationpolicy.org/programs/data-hub/international-migration-statistics>

### **Limitation Discussion - Raster Data Set 1**

For data set 1, the problem lies in the boxing of the year. This data set can only select time periods every 5 years, which is from 1950 – 1955 as the first category and 2015 – 2020 as the last category, instead of single years. Although the last year 2020 is explored in more detail in Data Sets 2 and 3, I did not find options to isolate other (1950 – 2019) single years in other data sets.

Some scholars have pointed out that the study of historical trends in immigration involves many complex public issues (Lapinski et al. 1997), if the time window selected is too broad, it is easy to lose details or make things complicated in the process. For example, if a researcher wants to study Japanese immigration and emigration during Japan’s bubble economy (1989 – 1991), then when the researcher uses the two-year boxes 1985 – 90, which refers to a 298,000 decrease and 1990 – 95, which refers to a 46,000 increase, other society events will be mixed in. Those social impactful events, such as the Plaza Accord in 1985 and the Great Hanshin earthquake in 1995, prevent researchers from being able to separate the variable of bubble economy well, and finally lead to a conclusion bias in research. In his paper, Tsuchida discussed the history and reasons for Japanese emigration, which showed a different motivation between different ages (Tsuchida 1998), which is powerful evidence of this argument.

### **Limitation Discussion – Vector Data Set 2 & 3**

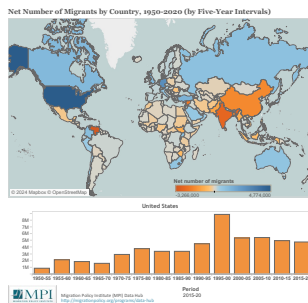
Data sets 2 and 3 together tell the story of immigrant inflows and outflows in 2020. For data set 2, readers can choose a country and then read where this country’s immigrants come from; for data set 3, shows the immigration and emigration of each country’s population in 2020. I hold the view that this part of the spatial data is on the one hand, too general and does not take into account the impact of world events, especially covid-19, and on the other hand, it is missing many demographic characteristics of immigrants.

The relationship between health and immigration is not a new issue, especially in the cross field to public health (Read et al. 2005, Kandula et al. 2004). This flaw will be further highlighted when scholars focus on immigration flows and COVID-19 public health (Page et al. 2020). The same possible shortcomings also occur in demographic characteristics. For example, in a study of farmers and workers using this data (Anderson et al. 2021), the researcher needs to merge the data sets based on the characteristics of occupation, which may cause incorrect variables. Problems such as different immigration counting methods lead to deviations in the final conclusion. In addition, more basic demographic characteristics such as gender and age, which also play a huge role in immigration research at specific time points (Berger 2013, Igoa 2013), are also missing in this data. This may lead a worse performance in the paper’s replication and reproduction. In addition, another very important flaw is that in both data set 2 and data set 3, the size of the circle only refers to the number of immigrants, instead of the immigrant population accounting for the total population of the country.

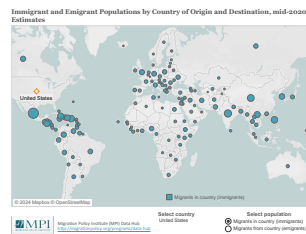
## Appendix

The Appendix shows the basic contents of the three data sets. Data set 1 selects the global view, data set 2 selects the united state as the base country, and data set 3 selects the immigrants view.

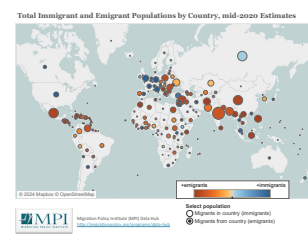
(a) Data set 1



(b) Data set 2



(c) Data set 3



## References

- Anderson, B., Poeschel, F. & Ruhs, M. (2021), ‘Rethinking labour migration: Covid-19, essential work, and systemic resilience’, *Comparative Migration Studies* **9**, 1–19.
- Berger, R. (2013), *Immigrant women tell their stories*, Routledge.
- Helbling, M. & Kalkum, D. (2018), ‘Migration policy trends in oecd countries’, *Journal of European Public Policy* **25**(12), 1779–1797.
- Igoa, C. (2013), *The inner world of the immigrant child*, Routledge.
- Kandula, N. R., Kersey, M. & Lurie, N. (2004), ‘Assuring the health of immigrants: what the leading health indicators tell us’, *Annu. Rev. Public Health* **25**, 357–376.
- Lapinski, J. S., Peltola, P., Shaw, G. & Yang, A. (1997), ‘Trends: Immigrants and immigration’, *The Public Opinion Quarterly* **61**(2), 356–383.
- Newland, K. & Patrick, E. (2004), ‘Beyond remittances: the role of diaspora in poverty reduction in their countries of origin, a scoping study by the migration policy institute for the department of international development’, *Migration Policy Institute* .
- Page, K. R., Venkataramani, M., Beyrer, C. & Polk, S. (2020), ‘Undocumented us immigrants and covid-19’, *New England Journal of Medicine* **382**(21), e62.
- Read, J. G., Amick, B. & Donato, K. M. (2005), ‘Arab immigrants: a new case for ethnicity and health?’, *Social Science & Medicine* **61**(1), 77–82.
- Stalker, P. (2002), ‘Migration trends and migration policy in europe’, *International migration* **40**(5), 151–179.
- Tsuchida, M. (1998), A history of japanese emigration from the 1860s to the 1990s, in ‘Temporary Workers or Future Citizens? Japanese and US Migration Policies’, Springer, pp. 77–119.