Patterns within Patterns: Using Python and Excel to Unlocking Hidden Migration and Settlement Patterns in Canada

Author: Francis Emmanuel Calingo

ABSTRACT

The vast majority of Canadians are either immigrants or descendants of immigrants, coming from a wide variety of countries and socioeconomic backgrounds. However, despite the importance of this matter, especially given the rapid increase of immigration in recent years (which will be shown later in this paper), there remains some gaps in both acquiring and publishing more granular data of the socioeconomic characteristics of the immigrant population, despite Statistics Canada's best efforts in doing so.

This project attempted to address some of those gaps by scraping census data and quarterly estimates from Statistics Canada over the past two decades, perform data cleaning and manipulation using Excel and Python, and visualize and analyze the data granularities and nuances of migration and settlement patterns, in order to present a national and subnational breakdown of the socioeconomic characteristics of the immigrant population (both permanent and non-permanent migration, and recent and non-recent immigrants). The results of this project yielded many interesting and useful insights. Firstly, the time-series analysis showed the rapid growth of the immigrant population, particularly non-permanent residents.Secondly, the categorical data analysis of recent immigrants (2016-2021) showed some discrepancies in settlement patterns, particularly in Quebec. The same can mostly be said with settlement patterns amongst the country's most populated census metropolitan areas, as well as the geospatial visualization of non-permanent settlement. These conclusions can help governments allocate their migrant resources to the right communities and implement more specialized migration policies.

Keywords: Migration, Settlement Patterns, Non-Permanent Residents, Socioeconomic classes

INTRODUCTION

In 1988, the Canadian Multiculturalism Act received Royal Assent, essentially rendering the preservation and promotion of Canada's ethnic diversity official government policy **[1]**. This passage epitomized the mainstream celebration and promotion of immigration in Canada. As per the 2021 Canadian Census, 8.3 million people in Canada (or nearly a quarter of the population) were born outside of Canada, coming from a wide variety of countries and socioeconomic backgrounds **[2]**. Settlement patterns in Canada in recent decades have been concentrated in the country's largest metropolitan areas, especially the Greater Toronto Area, Metro Vancouver, Greater Montreal, and the Calgary Metropolitan Area. However, we are starting to witness more migrants settling in places that have not been associated with . For instance, in the 2006 Census, less than 4% of residents of Prince Edward Island were born outside of Canada **[3]**, but by 2021, that percentage has ballooned to 11.4% **[4]**.

This kind of diversity not only renders data analysis of migration and settlement patterns in Canada a very interesting research topic of exploration, but also renders it highly important for helping guide policymakers with regards to helping migrants of varying backgrounds integrate in Canadian society. For example, a community with a high proportion of non-permanent residents that come from a country where neither French nor English is widely spoken nor understood will need a lot of federal government support and funding in helping their non-permanent resident community adapt to their new home. Despite the aforementioned gaps with regards to more granular research on the socioeconomic characteristics of settlement patterns, more resources must not only be allocated towards this research topic, but must also be more robust. In doing so, this shift will help governments and civic organizations unlock hidden patterns that may have not been so apparent such as a gendered breakdown of migrants by their countries of origin. Just like how race-based data collection and analysis were key in unlocking racial inequalities during the height of the COVID-19 pandemic **[5]**, more granular research on migration and settlement patterns can potentially help unlock hidden socioeconomic inequities that are not always known nor discussed in mainstream migration discourse.

[1] Full text of the law: https://laws-lois.justice.gc.ca/eng/acts/C-18.7/FullText.html

[2] Government of Canada, Statistics Canada. *The Daily — Immigrants Make up the Largest Share of the Population in over 150 Years and Continue to Shape Who We Are as Canadians.* 26 Oct. 2022, https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm.

[3] Government of Canada, Statistics Canada. *Statistics Canada: Visual Census.* 7 Dec. 2010, https://www12.statcan.gc.ca/census-recensement/2006/dp-pd/fs-fi/index.cfm?LANG=ENG&VIEW=D&PRCODE=11&TOPIC_ID=6&format=jpg.

[4] Government of Canada, Statistics Canada. *Focus on Geography Series, 2021 Census - Prince Edward Island.* 13 July 2022, https://www12.statcan.gc.ca/census-recensement/2021/as-sa/fogs-spg/page.cfm?lang=E&topic=9&dguid=2021A000211.

[5] An example of the usefulness of race and income-based data collection is Toronto Public Health's dashboard on the unequal effects of COVID-19 on Toronto's racialized and low-income communities (although it is inactive as of January 2022): https://www.toronto.ca/community-people/health-wellness-care/health-programs-advice/respiratory-viruses/covid-19/covid-19-pandemic-data/covid-19-archived-dashboards/covid-19-ethno-racial-identity-income/

This is not to say that this shift has not started already. In fact, as migration and settlement patterns continue to diversify, we are starting to witness a shift in migration literature that focuses less on migration as a whole and more on specific communities and specific variables of migration. For instance, Statistics Canada, in collaboration with the Office of the Special Representative on Combating Islamophobia, recently published a publication visualizing the demographics of the Canadian Muslim population, including settlement patterns of migrants from Muslim-majority countries **[6]**. Publications like that one are uncommon for the time being, but they should be the norm eventually.

DATA COLLECTION

Datasets

The following links will redirect you to each census' data on recent immigrants **[7]** since 2001:

2001 Census Topic-based tabulations:
https://www12.statcan.gc.ca/english/census01/products/standard/themes/Rp-eng.cfm ?LANG=E&APATH=3&DETAIL=1&DIM=0&FL=A&FREE=0&GC=0&GID=0&GK=0&GRP= 1&PID=62124&PRID=0&PTYPE=55430,53293,55440,55496,71090&S=0&SHOWALL=0&S UB=0&Temporal=2001&THEME=43&VID=0&VNAMEE=&VNAMEF=

Place of birth for the immigrant population by period of immigration, 2006 counts and percentage distribution, for Canada, provinces and territories - 20% sample data:
https://www12.statcan.gc.ca/census-recensement/2006/dp-pd/hlt/97-557/T404-eng.cfm?Lan g=E&T=404&GH=4&GF=1&SC=1&S=1&O=D#FN2

2011 National Household Survey: Data tables: Place of Birth, Period of immigration=2006-2011: [8]
https://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/dt-td/Rp-eng.cfm?TABID=2&LANG=E&AP ATH=3&DETAIL=0&DIM=0&FL=A&FREE=0&GC=0&GK=0&GRP=1&PID=105411&PRID=0 &PTYPE=105277&S=0&SHOWALL=0&SUB=0&Temporal=2013&THEME=95&VID=0&VNA MEE=&VNAMEF=

[6] Government of Canada, Statistics Canada. *The Muslim Population in Canada.* 16 Dec. 2024, https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2024058-eng.htm.

[7] For their censuses, Statistics Canada defines "recent immigrants" as people who have migrated to and settled in Canada within the last 5 years of the census enumeration date (i.e., between two consecutive census enumeration dates, exclusive of the previous census enumeration date).

[8] The mandatory long-form census was scrapped for 2011 (being replaced by the voluntary National Household Survey) before being reinstated in 2016. Therefore, there may be some data anomalies, although their statistical significance is unknown for the scope of this project. Regardless, for convenience, the 2011 National Household Survey will be referred to as a "census" for the duration of this paper.

Census Profile, 2016 Census: Recent immigrants by selected places of birth:
https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=PR&Code1=01&Geo2=PR&Code2=01&SearchText=Canada&SearchType=Begins&SearchPR=01&B1=Immigration%20and%20citizenship&TABID=1&type=0

Census Profile, 2021 Census of Population: Selected places of birth for the recent immigrant population:
https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/page.cfm?LANG=E&GENDERlist=1,2,3&STATISTIClist=1,4&DGUIDlist=2021A000011124&HEADERlist=26&SearchText=canada

The following links will redirect you to Statistics Canada quarterly data on migration:

[Archived - Estimates of non-permanent residents, quarterly, inactive, 2001-2018]
https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710002301&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2001&cubeTimeFrame.endMonth=10&cubeTimeFrame.endYear=2018&referencePeriods=20010101%2C20181001

[Estimates of the components of international migration, quarterly, 2021-2024]
https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710004001&pickMembers%5B0%5D=1.1&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2018&cubeTimeFrame.endMonth=10&cubeTimeFrame.endYear=2022&referencePeriods=20180101%2C20221001

[Estimates of the number of non-permanent residents by type, quarterly]
https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710012101&pickMembers%5B0%5D=1.1&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2021&cubeTimeFrame.endMonth=07&cubeTimeFrame.endYear=2024&referencePeriods=20210101%2C20240701

Data Limitations

There is no available data for the total number of non-permanent residents in Canada between Q3 2018 to Q2 2021. Therefore, the quarterly net non-permanent migration from that timespan was used to calculate for the missing values.

Additionally, keep in mind that some variables of the 2021 Census used a 25% sample size, especially with data on recent immigrants, while other aspects such as total population enumerates the entire population.

METHODOLOGY

The analysis will be broken down into four parts: (1) Time-series analysis of international migration to Canada over the past two decades, (2) Categorical data analysis of the socioeconomic characteristics of recent immigrants as per the 2021 Census, (3) Scatter plot analysis of migration patterns for Canada's Top 25 [9] census metropolitan areas (CMAs) by population as per the 2021 Census, (4) Geospatial analysis of recent non-permanent migration. Except for the geospatial plot, every other plot will call on the Plotly and Seaborn libraries of Python.

Time-series analysis: The first two plots are area charts (subdivided by provinces and territories). The first plot visualizes the cumulative total of non-permanent residents (quarterly since Q1 2001), and the second plot is net quarterly international migration.
The third plot will take the top 10 countries of origins for people that migrated to Canada in 2020 [10], taking each of their estimated total from the last 5 census' "recent immigrants" data (2001, 2006, 2011, 2016, 2021), and plot each in a line graph.

Categorical data analysis (2021 Census data): The first plot will be a bar plot for each province and territory. Each bar represents the % of recent migrants by country of origin in province and territory that identify as woman, and this dash line represents the percentage of the general population within each geographic area that identify as a woman.

The next plot will be a stacked bar chart that visualizes each province and territory's top 5 countries of origin amongst recent migrants.

The third plot will be a grouped bar graph that compares the age distribution of both the general population and the foreign born population of each province and territory. The age groupings differ for both groups:

Total - Age groups of the population - 100% data	Total - Age at immigration for the immigrant population in private households - 25% sample data [84]
0 to 14 years	
0 to 4 years	
5 to 9 years	
10 to 14 years	
15 to 64 years	Under 5 years
15 to 19 years	
20 to 24 years	
25 to 29 years	5 to 14 years
30 to 34 years	
35 to 39 years	
40 to 44 years	15 to 24 years
45 to 49 years	
50 to 54 years	
55 to 59 years	25 to 44 years
60 to 64 years	
65 years and over	
65 to 69 years	45 years and over
70 to 74 years	
75 to 79 years	
80 to 84 years	
85 years and over	

[9] For the purpose of this analysis, the Ottawa-Gatineau CMA was broken up into two parts: its Ontario and Quebec part, rendering the analysis of the "Top 25" CMAs Top 26.

[10] *ARCHIVED – 2021 Annual Report to Parliament on Immigration*, Immigration, Refugees and Citizenship Canada, n.d., www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/annual-report-parliament-immigration-2021.html

The left image shows a more granular breakdown of age distribution for the general population, while it is more generalized for the immigrant population. Therefore, arithmetic operations were executed on the age distributions of the general population so that data for 5 specific age ranges (under 5, 5-14, 15-24, 25-44, 45+) could be ascertained.

Scatter plot analysis: The first plot will compare CMA population to each CMA's non-permanent resident population per capita (per 100,000), with points categorized by region (i.e., BC, Prairies, Ontario, Québéc, Atlantic). Both populations are based on the 2021 Census enumerations. The per-capita non-permanent resident population were calculated using this formula:

$$\frac{[2021\ Non\ Permanent\ resident\ population]}{[2021\ CMA\ population]} * \frac{1}{100,000}$$

Secondly, the growth rates of the CMAs' general population and their immigrant population (2016-21) were compared. In both cases, their growth rates were calculated using this formula:

$$\frac{([2021\ census\ population] - [2016\ census\ population])}{[2016\ population]} * 100$$

The third plot will be a series of scatter plots comparing CMA populations to per-capita (per 100,000) recent immigrant populations broken down by continents of origin, and colours once again assigned by region.

Geospatial analysis: The growth rate of the non-permanent resident population between Q3 2022 and Q3 2024 will be mapped using the Geopandas library of Python. Similar to a formula above, the growth rate is calculated by:
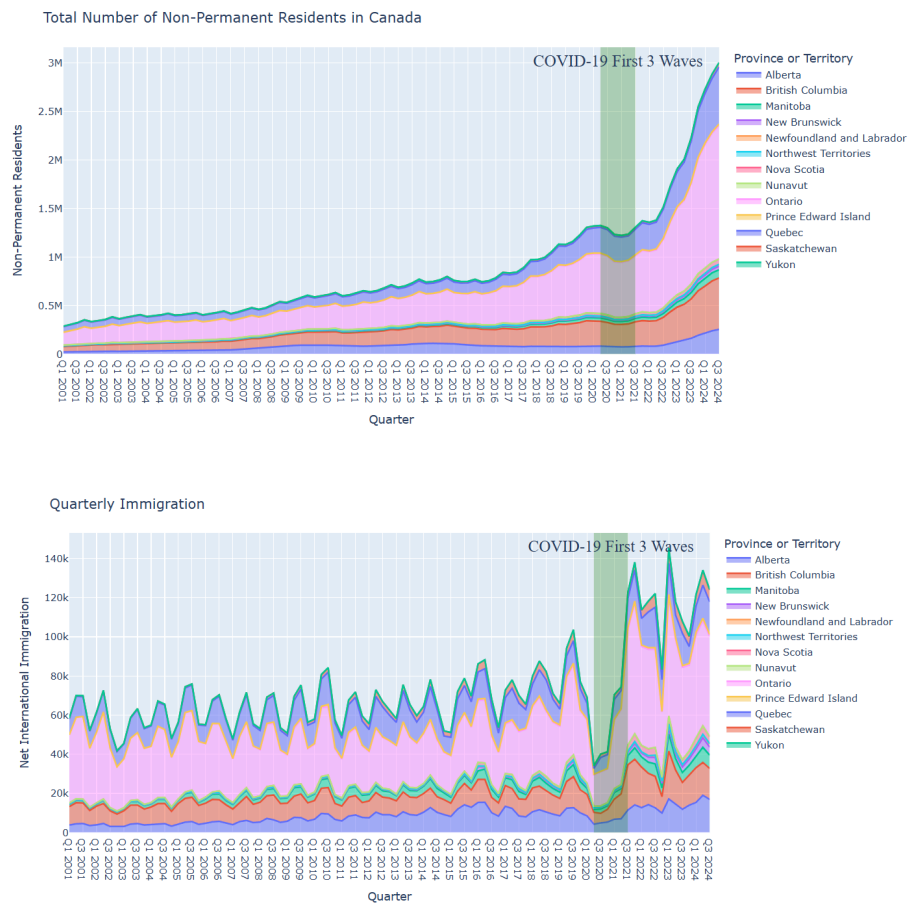
$$\frac{([Q3\ 2024\ Non\ Permanent\ resident\ population] - [Q3\ 2022\ Non\ Permanent\ resident\ population])}{[Q3\ 2022\ Non\ Permanent\ resident\ population]} * 100$$

RESULTS

Time-series analysis:

The first plot is a breakdown of the number of non-permanent residents in Canada quarterly by provinces and territories, you'll notice the dip in NPRs during the first three waves of COVID-19, reflecting the travel restrictions of the time, as well as the accelerated growth after 2022, reflecting domestic labour shortages.

The next plot is net quarterly international migration. Once again you see the effects COVID had on migration, as well as the increase in numbers after all COVID-related restrictions were lifted. You can also see the seasonality of migration with all these peaks and valleys.


Total Number of Non-Permanent Residents in Canada

Quarterly Immigration

An advantage of Plotly is its interactivity. In this case, it's the ability to look at province(s) and/or territory (or territories) in isolation. For example, let's examine Nova Scotia's non-permanent resident population and Saskatchewan's quarterly net international migration:

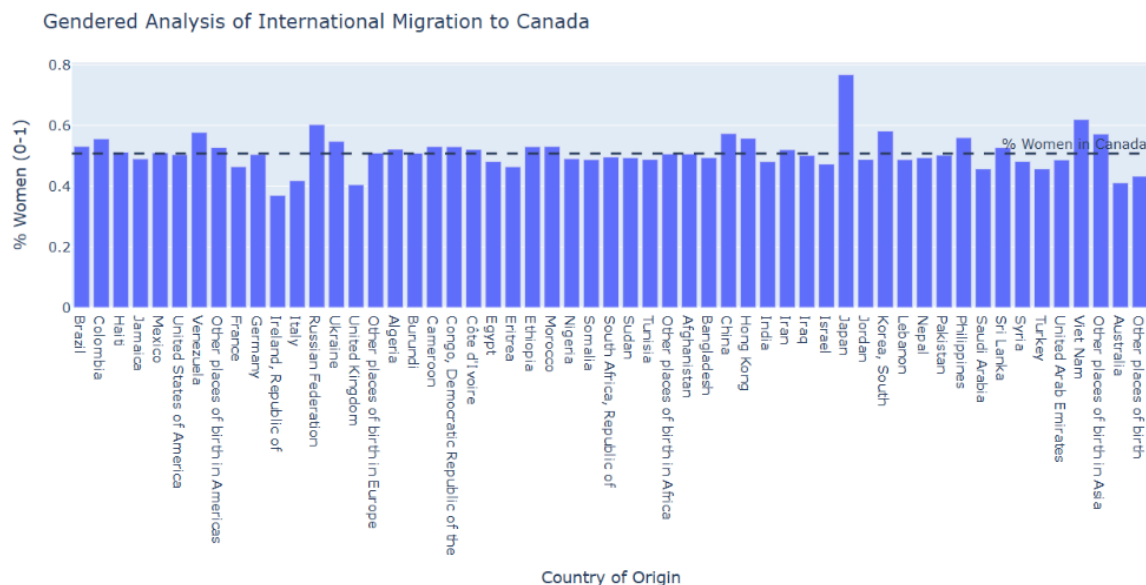Total Number of Non-Permanent Residents in Canada
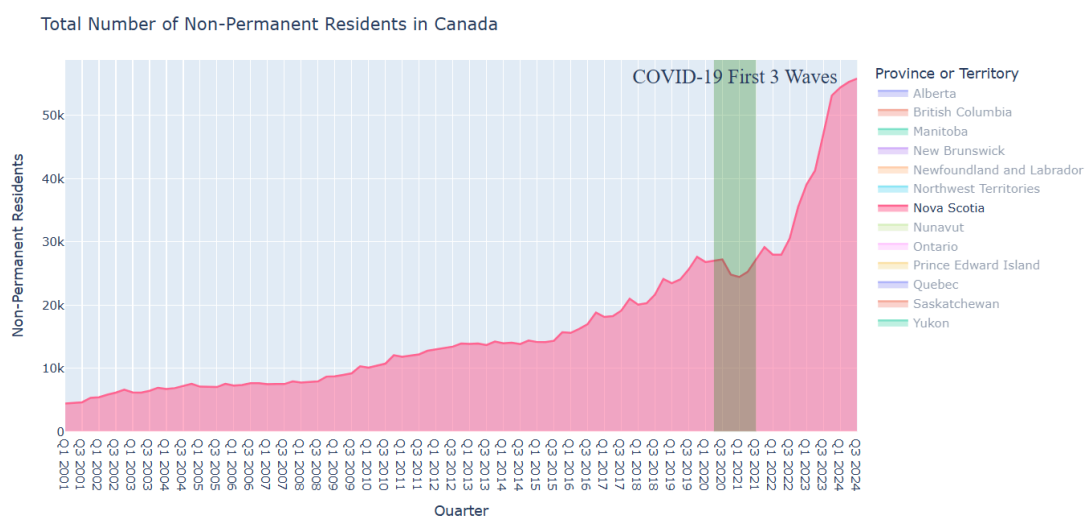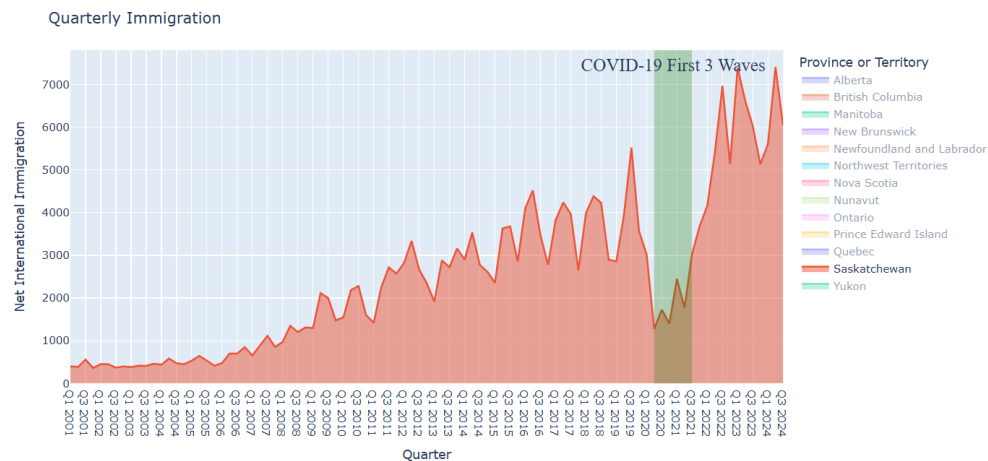
The next plot breaks migration patterns of recent migrants down census-by-census, using the top 10 countries of origin from 2020. China, India, and the Philippines have consistently been Canada's top 3 countries of origin, with recent immigrants from India especially seeing an accelerated growth.

Observe the rapid growth of the Syrian migrant population in the 2010s coinciding with the Syrian Civil War as well as the more open-door migration policy since the Justin Trudeau government was voted into power in 2015.
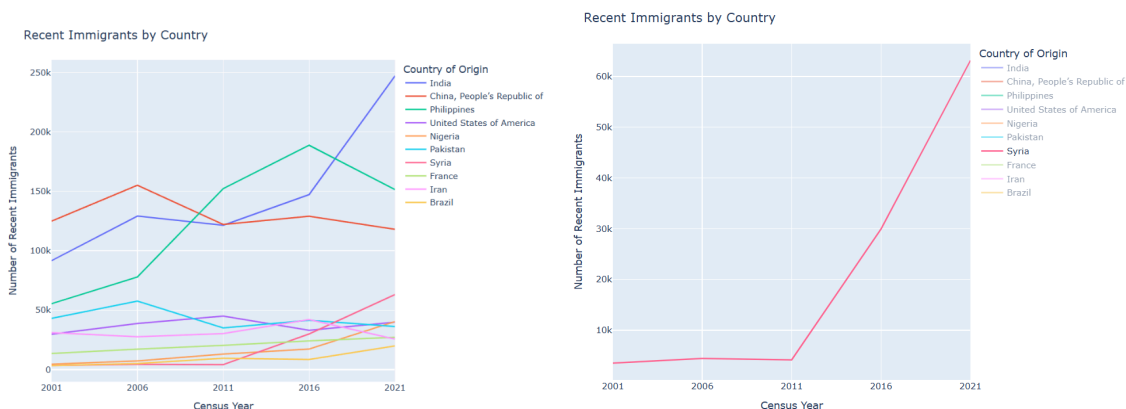

Categorical data analysis:

The first plot is a gender analysis of recent migrants (2016-21) by province and territories. Each bar represents the % of recent migrants by country of origin in province and territory that identify as woman, and this dash line represents the percentage of the general population within each geographic area. Observe that there are some variations by country, on one extreme is Japan with the largest positive percent difference, and on the other extreme is the Republic of Ireland. If we look at the different provinces and territories, some countries of origin display consistent behaviour like Japan, and others like Burundi (in Alberta), we'll observe that there is a much bigger difference.

Gendered Analysis of International Migration to Canada

Gendered Analysis of International Migration to Alberta
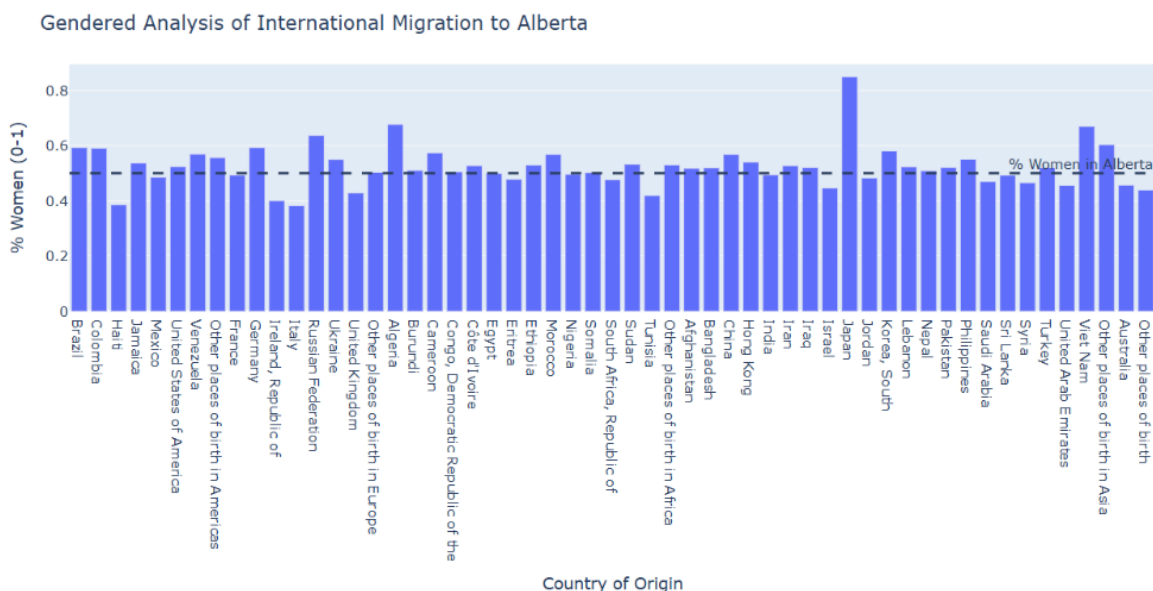
The next plot is a stacked bar chart that looks at each province and territory's top 5 countries of origin amongst recent migrants (2016-21). Some countries like China consistently show up. There are a few interesting cases such as the US and UK making it to British Columbia's Top 5. Québéc's migration pattern stood out, where even its linguistic culture has an impact on migration, as all but one of the top 5 countries are francophone countries or countries where French is widely understood.

Breakdown of Country of Origin by Province/Territory
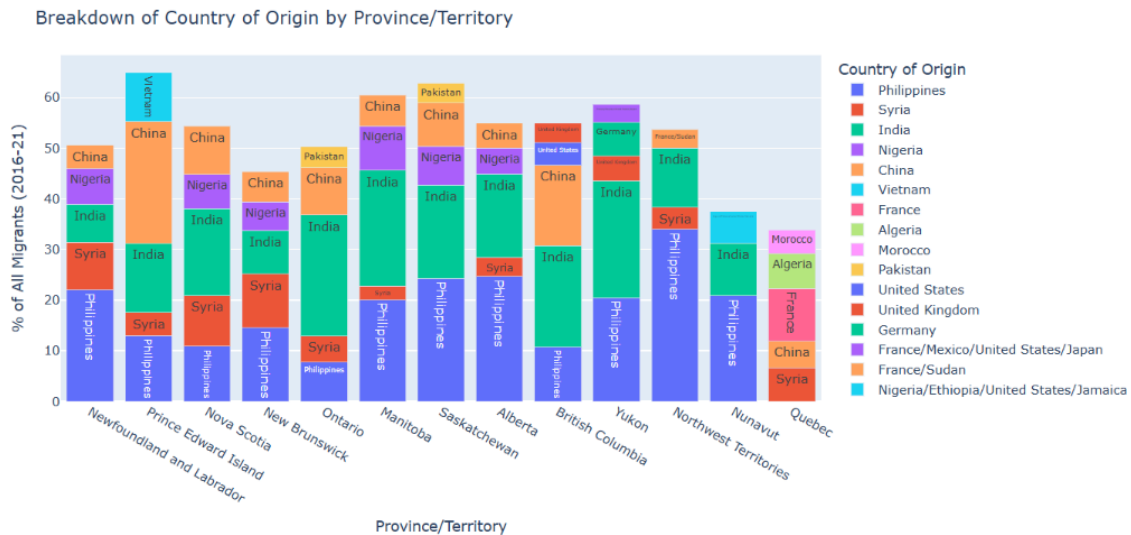
As this is a Plotly plot, we can see certain aspects of the plot in isolation. For example, when we filter for the Philippines, we can see that recent immigrants from the Philippines take up a significant portion of the recent migrant population in places such as Newfoundland and Labrador and NWT (suggesting that there is a vigorous campaign to have Filipinos migrate there). Conversely, they don't even make the Top 5 In Québéc (likely due to language barriers).

Breakdown of Country of Origin by Province/Territory
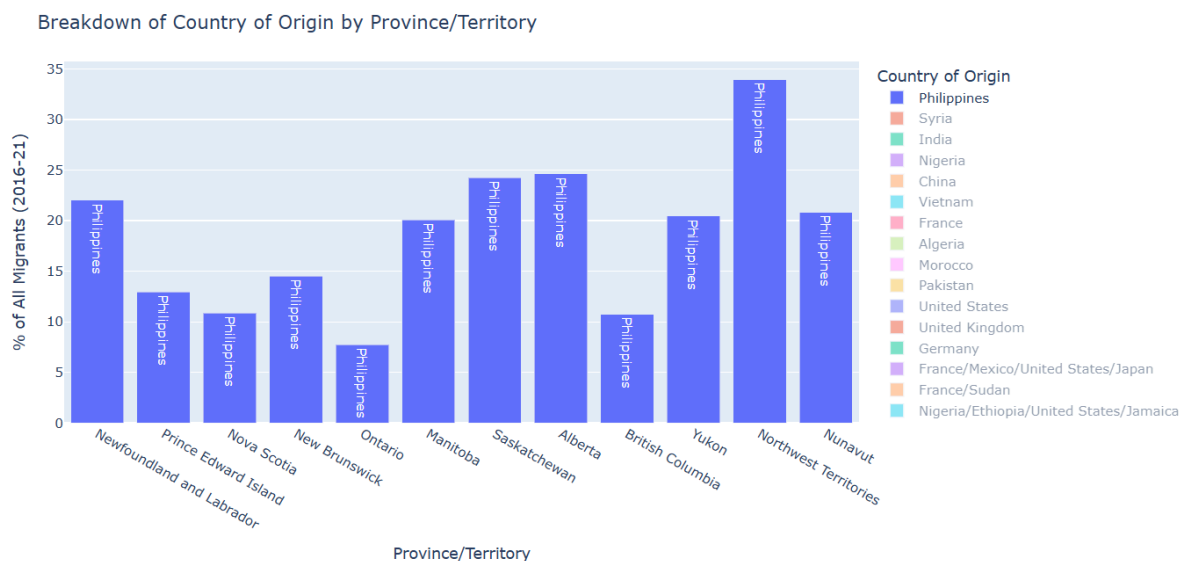
The next plot is an age-based analysis comparing the age distribution of the general population versus the total immigrant population, again using 2021 census data. This FacetGrid shows a general pattern that people aged 25-44 make up the plurality of the foreign-born population, but those aged 45+ make up the plurality of the general population. Of course, there are a few interesting cases such as Nunavut, where the young age on average as a whole is visualized well here. So it appears that the characteristic of the foreign-born population is more oriented towards working-age adults than the general population.
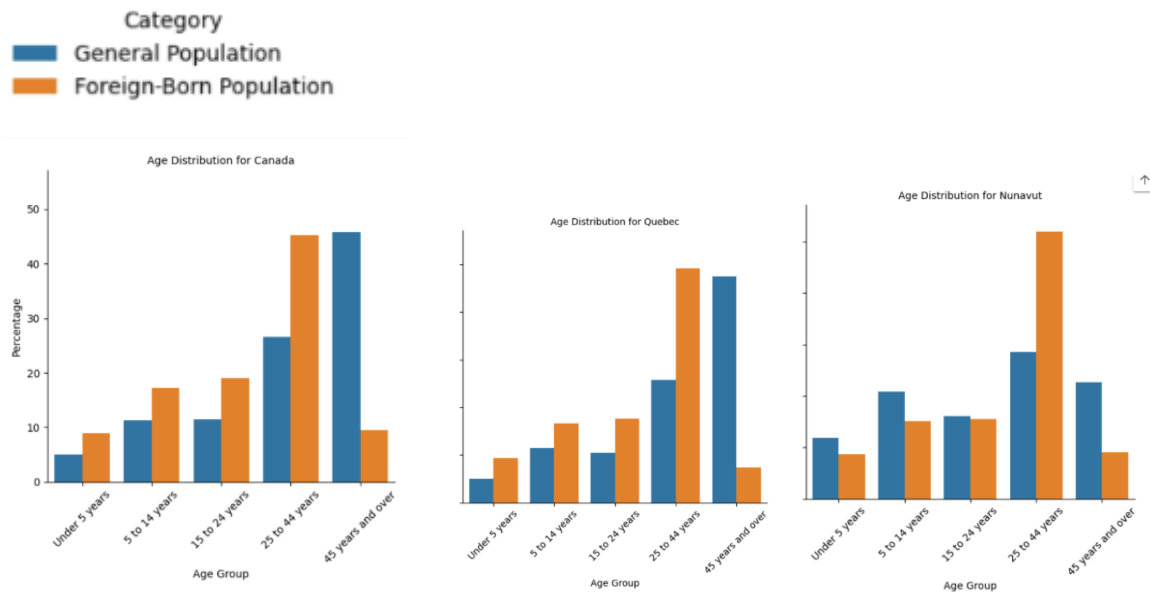
Scatter plot analysis:

This first scatter plot compares CMA population to each CMA's non-permanent resident population per capita, with points categorized by region (i.e., BC, Prairies, Ontario, Québéc, Atlantic). You can see that outside of Montréal, Québéc has quite a low per-capita non-permanent resident population. Everything else is pretty much a mixed bag. How much of this discrepancy observed in Québéc is due to linguistic barriers and/or societal attitudes towards immigrants is not known.


Secondly, the growth rates of these CMAs and their immigrant population (2016-21) were compared. Again they are broken down by region. And based on this graph, you can see that pretty much every CMA's foreign-born population is outpacing their general population's growth rate, with a bit of regional variation. BC exhibited the highest general population growth rate, but the Atlantic exhibited the highest foreign-born population growth rate, possibly due to both the federal and provincial governments' push to attract more immigrants to the Atlantic region in a bid to offset population loss and inter-provincial out migration.

Growth Rates of CMAs vs. CMAs' Foreign-Born Population
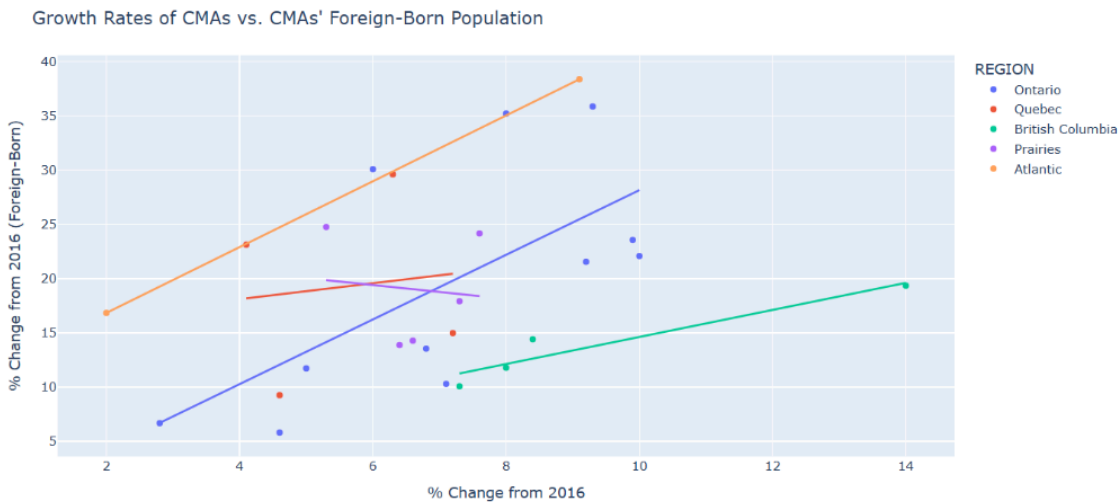
The third plot is a FacetGrid of some scatter plots comparing CMA populations to per-capita recent immigrant populations broken down by continents of origin, and colours once again assigned by region. Note that using Seaborn instead of Plotly changed the colouring scheme. For the most part, there were no interesting nor discernible patterns for the rate of migration for people from the Americas, Europe, and Oceania. But where it gets interesting is looking at the plots for Africa and Asia. It's more dispersed for the Asia plot versus the Africa plot.
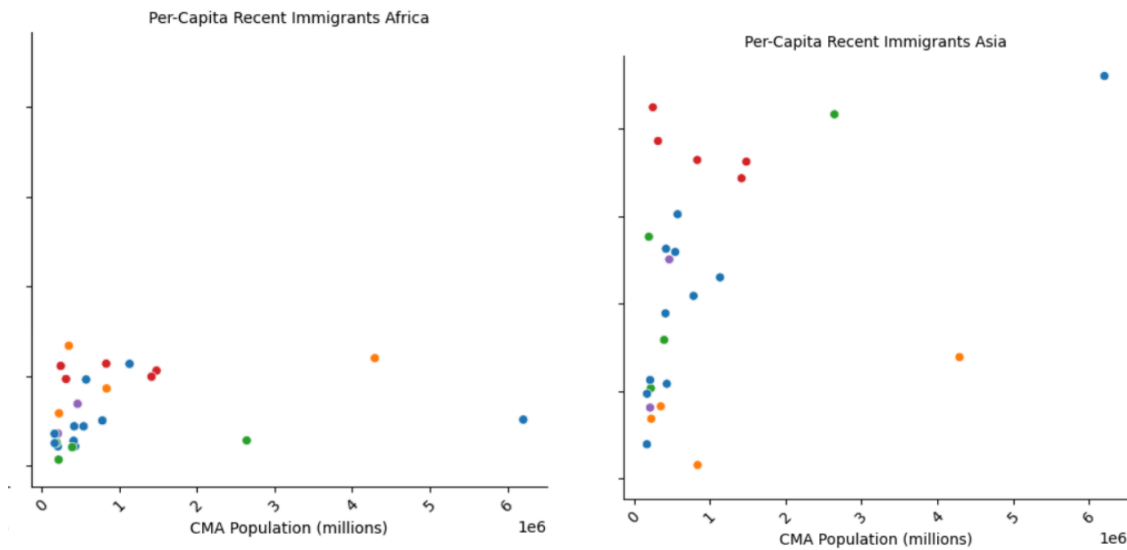
Observe the point circled in red below. This is Montréal, observe the difference between both plots. It appears that migrants of African origin, on a per-capita basis, is comparatively high compared to other CMAs, while it is comparatively lower with migrants of Asian origin. And to some extent it rings true for the rest of Québéc.
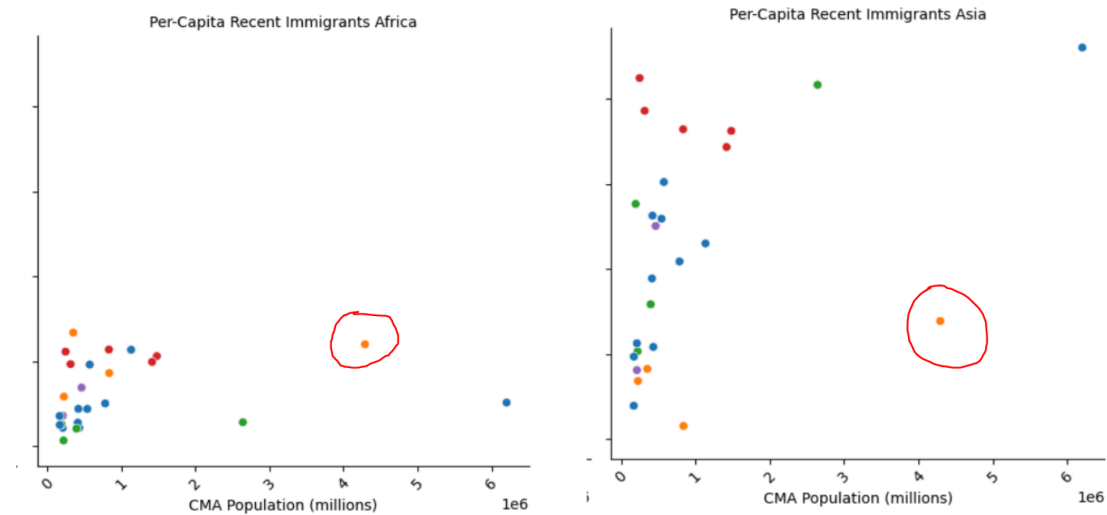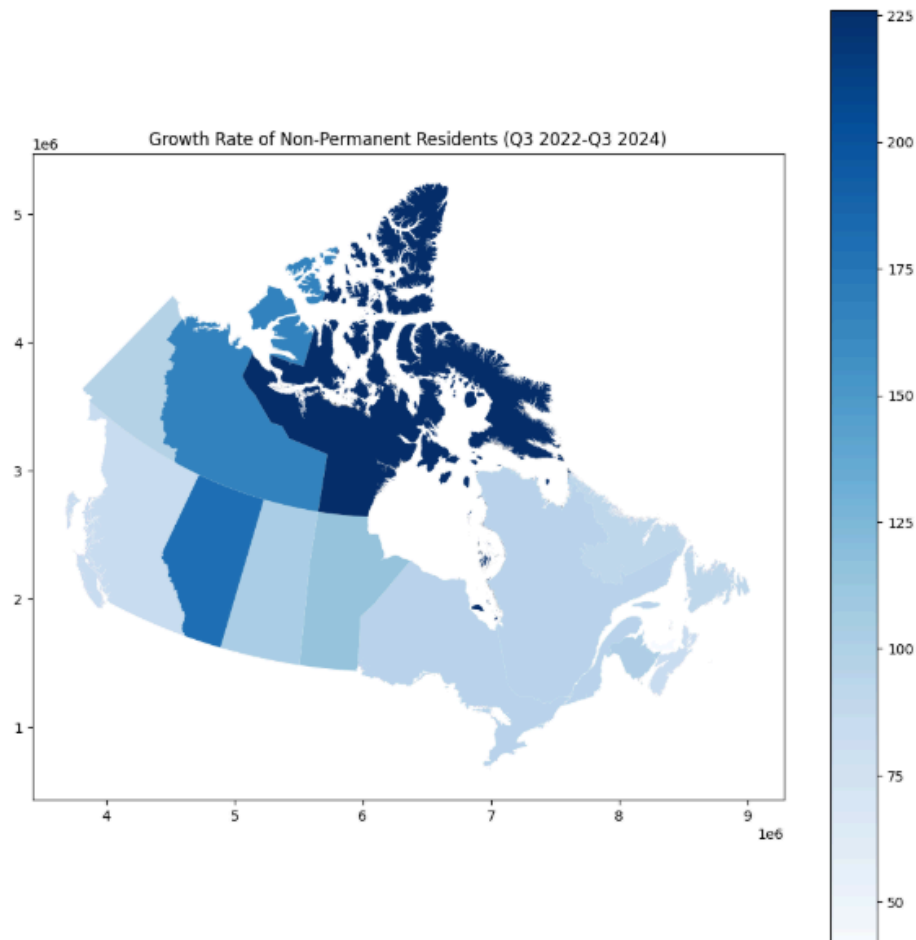

Also, the cluster of red points below represents Prairie CMAs, suggesting that they share a similar migration pattern.
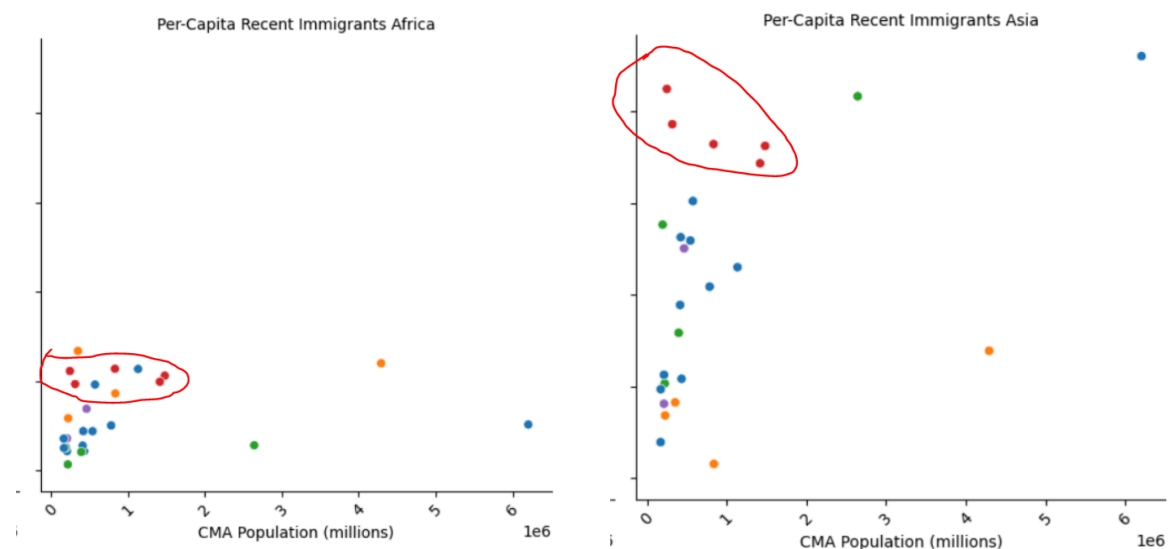
Geospatial analysis:

Recall that the national non-permanent resident population rose dramatically after Q3 2022. Here is a choropleth map showing a geographic breakdown of that growth.

Growth Rate of Non-Permanent Residents (Q3 2022-Q3 2024)

Nunavut by far experienced the fastest non-permanent resident growth within that time period, although that can largely be attributed to its small non-permanent resident. Alberta also experienced high growth, mirroring its generally high growth as a whole over the past decade. Growth has slowed down for British Columbia and Ontario. Surprisingly, the Maritime provinces experience relatively little growth despite government initiatives to attract immigrants and foreign labour into that region.

CONCLUSION

There were definitely some interesting insights and surprises, notably the gendered analysis. It would be interesting to know what is or are the root causes of the gender imbalance of recent migrants from Japan. Perhaps there is a labour component to this imbalance in the sense that most recent immigrants from Japan work in a certain sector that could be heavily gendered such as healthcare.

By making this kind of data more granular and adding a spatial and geographic component, policy decisions both within and outside Canada can be remolded to better cater to communities with different needs and socioeconomic conditions. For example,

- A country's consular services can be redirected more towards underserved communities, especially communities with high levels of non-permanent migration such as Nunavut.
- Determine which local and provincial/territorial governments will have to do more to provide more specialized services to help certain communities integrate well.
- Increase more women-specific migrant resources for certain migrant communities such as the Burundi diaspora in Alberta.
- Examine Québéc's unique migration patterns, and explore the different ways its language laws have an effect on them.