



VISUAL INVENTORY RESEARCH PROJECT

University of Ottawa
EBC 6997

Yu Mi
300041818

Introduction

Canada offers a myriad number of services to its newcomers, including but not limited to language training, job search, housing services, and etc. Both the federal and provincial governments have been funding these services to immigrants in hopes of helping them establish a new home in Canada. The effectiveness of an additional service provider depends on the location and types of services its close vicinity already offers and existing or future demands in a local area. From the newcomer perspective, a geographical visual of service provider locations can provide them with a one-stop-shop while searching for local options. Having access to a visual database will enable service providers to better understand and adapt to each community's needs. They will be able to establish themselves in places with high service level gaps. The intent of this project is to create a web-enabled and interactive prototype of newcomer service provider inventory and recommend strategies for continued data collection, database update, and technological modifications.

Several data visualization examples were consulted prior to the start of the project. One example from the online journal [CarbonBrief](#) used Leaflet JavaScript library to create a map using Mapbox as the background map layer. This visual contains information regarding United States' electricity generation based on state, type, and capacity. Colour and size of symbols were used to demonstrate capacity size and type of electricity. Users also have the option of viewing the visual in satellite, daytime, or greyscale imagery. The article draws several conclusions based on the geographical placement of different energy plant types. In addition, [Quebec's immigration service provider interactive map](#) as also consulted. This map shows a single symbol for all service types, each service provider's contact information can be found on the left-side panel upon left-click. However, this map does not distinguish different types of services from one another.

Scope and Definition

This project intends to address the concerns of all three stakeholders: government, newcomer, and service provider. Service providers include organizations which are providers of settlement and integration services, not simply locations hosting the services. Newcomer is defined as anyone who has come to Canada for a long-term stay (longer than one year) within the last 5 years from the date the data was gathered. This is under the assumption that those who have resided in Canada for longer than 5 years will no longer require newcomer services. Moreover, it is also the eligible duration for service provider to receive government grants.

Methodology

This project's datasets come from various sources. The federal repository of newcomer service providers came from the [IRCC \(Immigration, Refugees, and Citizenship Canada\) website](#) and was downloaded as a CSV file for further processing. This dataset contains contact information, geolocation information (longitude and latitude), and divides newcomer services into 13 categories: job search, women, seniors, youth, language training, language assessment, job-specific language training, daily life, francophone services, LGBTQ2, become a mentor, and

others. [Quebec's repository](#) is standalone. Unlike the federal repository, this province has identified 28 types of services. In order to display a consistent national database, these 28 types are mapped into federal's corresponding 13 types as per Appendix A. Canadian population and immigration statistics were extracted from various locations on the Statistics Canada website.

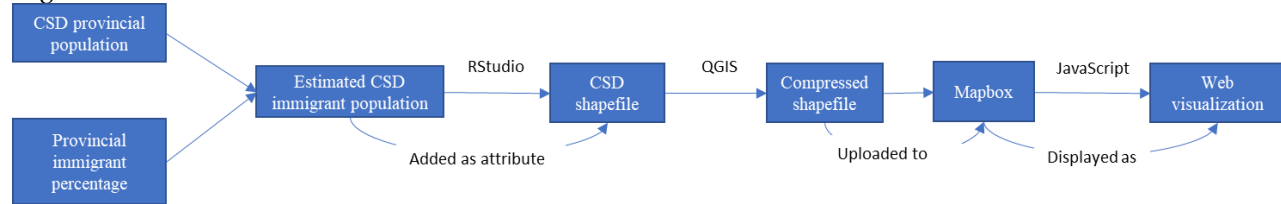
The visual representation of newcomer service providers is achieved by uploading their associated CSV files into Mapbox (an online platform to represent locational data). The CSV files became datasets which then were transformed into tilesets and styles. The service provider's contact information is shown as a popup when clicking on the dot. Each type of service is uploaded as a separate layer from the available styles and represented by a uniquely coloured dot. Each layer can be turned on or off by clicking on the layer menu on the right-hand side of the map. Mapbox GL JS JavaScript library was used to bring the map from Mapbox onto a website (<http://kevinkells.com/map/>). For ease of collaboration and accessibility, this JavaScript is maintained on [Github](#).

After creating the online inventory in Mapbox, this prototype was shown to two newcomers and one prior IRCC employee for feedback. Improvements were made after the feedback sessions to address user concerns. User concerns were recorded as issues in GitHub and resolved through pull requests. Notably, a checkmark now appears whenever a layer is turned on and driving directions has been added.

One of the major enhancements added as a result of the feedback sessions is a predicted immigrant population map layer. The choice of the heat map colour scale is eight pseudo-logarithmic steps plus zero: 0, 10, 50, 100, 500, 1,000, 5,000, 10,000, and 50,000. The estimated immigrant population was derived from multiplying [2016 Census Subdivision \(CSD\) provincial population](#) by provincial immigration percentage (calculated by dividing [number of 2011-2016 immigrants by total provincial population](#)). RStudio was used to extract relevant records from the CSV files containing CSD population information and perform the calculation. It would have been extremely tedious if this process were performed in Excel due to the files' size and number of records. In some cases, Excel could not even read all of the information contained in document.¹ Then, using the same software, this information was consolidated in a CSV file and added as an attribute to the CSD shapefile provided by the [Statistics Canada website](#). Finally, the shapefile was compressed in QGIS (a geographic information system software) by 100 meters of tolerance so that the data became small enough in size to be uploaded into Mapbox (or else there will always be a system timeout). The size of the compressed folder was reduced from 47.8MB to 9.5MB from this process. Even if Mapbox successfully uploads a large size tileset, it will automatically adjust the zoom level; disabling the layer's visibility until the map is zoomed in to a certain level. This means the layer will not be visible when the map is zoomed out to show the entire country or a specific province or territory. Figure 1 below is a diagram of the process and technologies used for the data transformation.

¹ As mentioned below in the section "Heat Map Layer," the assumed similarity between newcomer distribution and general population distribution was an initial approximation whose validation went beyond the scope of the present project.

Figure 1



Results

In general, newcomer services exist in locations which are predicted to be densely populated by immigrants. The number of estimated immigrants range from 0 to 95,878: 8.18% of the CSDs are 0 and 2.89% of the CSDs are over 1000 people. This suggests newcomers are expected to be highly concentrated in only a few CSDs. There are also 14 CSD without population data. Large cities such as Toronto, Montreal, and Vancouver all have high concentrations. The map also shows services exist in smaller, remote cities such as Whitehorse, Iqaluit, Yellowknife, and Chibougamau. Hence, the existing service providers have already been somewhat strategically placed. On the other hand, one can only visually inspect the data points against population concentration. There is no quantitative comparison against supply of services and quantity demanded.

It should be noted that there are several regions with unexpectedly high population concentrations such as northeast Alberta, Greater Sudbury, and the border between Ontario and Quebec. Alberta has the highest provincial recent immigrant percentage of 5.11%, and its oil fields are located in the north, so a higher estimated newcomer population there would make sense, if validated. However, there are only a handful of services (approximately eight) near Fort McMurray, suggesting may be a shortage of service providers in the region. Similarly, Rouyn-Noranda, Greater Sudbury and Timmins are somewhat densely populated cities (estimated around 1,000 newcomers based on our approximation) heavily dependent on natural resource industries. These cities also only have a limited number of service providers (approximately eight).

Cities such as Kenora, ON and Sydney, NS are estimated to have 1000 immigrants in certain CSDs but unfortunately newcomer services do not exist at all in these locations. Moreover, it should be noted west and central Ontario has a considerable number of residents nonetheless its newcomer service providers are almost non-existent.

Discussion

Data Collection and Maintenance

This project first began with the intent of analyzing a single city, Winnipeg; then it expanded into the entire country when a public national database was discovered from the IRCC website. This allows for easy maintenance of a national-wide dataset. One downside of using this

readily available dataset is research is limited to the records available from the database. For example, this database does not identify community centres or nearby high schools for immigrants with children. In addition, it is suspected that this dataset only accounts for organizations who have received funding from the government (they are distinguished into categories at time of successful funding allocation); applicants are required to identify the service(s) provided to newcomers as a part of their funding request process. Therefore, this repository will not encompass non-registered self-funded service providers (for example those who provide their services in exchange for a fee). Compared to the manual research completed for Winnipeg, Manitoba, the newcomer services identified under the IRCC website did not encompass all services found for this city. This suggests not all newcomer services have received funding from or are registered with the Government of Canada. A way to improve IRCC's current database is to request businesses to self-identify as a service provider at time of incorporation. In addition, an annual survey could be conducted with business owners to verify their service provisions.

One observation made during the source data mapping process is that its dataset nomenclature does not exactly match that of the IRCC website. For example, "refugee" category on the website became "help GAR (Government Assisted Refugee)", "become a mentor" became "volunteer", and "help with daily life" became "general". These results were confirmed by matching website results under these designated categories with those from the source document. The reason for the discrepancy is the database developers who created the database did not use the same naming convention as the technicians who created the website. This is a result of the difference in target audience. For the database creators, their main stakeholder is intra-departmental, therefore their naming convention is IRCC centric (using the department's jargon). On the other hand, the website developers' target audience is external stakeholders, which means the category names need to be clear and easy to understand to the general public. It is recommended to establish a consistent nomenclature to avoid confusion between users and developers.

The recommended time interval to update the national service provider inventory database is annual because funding is allocated on an annual basis after the call for proposal and assessment of application approval process. There is no requirement to refresh the datasets more frequently than once per year. Projects are signed in March for commencement in April (start of government fiscal year) therefore dataset update can be completed in May in order to compensate for possible delay in government's dataset upload onto the website. If there are self-identified service providers who don't exist in the IRCC website, then their information including geocoded locations can be added as new records. Geocoding can be accomplished using either python or R. As for CSD census data and shapefile, they can be refreshed every 5 years as per Statistics Canada's time schedule. The next set of national census data for 2017 – 2022 will be available in 2023. The same R script can be used to calculate and add estimated immigrant population to the shapefile (the next section suggests some minor improvements to the R script).

Technology

This project was initially planned to be visualized in Tableau because Tableau is a mature data visualization software capable of representing location data and it is also widely adopted by both private and public industries. It is often the standard for operational and strategic dashboards in large organizations. Unfortunately, its geographic information processing is limited, and it was impossible to show more than two mutually inclusive location points in a simple, clean, easy-to-read format. It was determined that Mapbox would be the most appropriate platform for this project after consulting several cartographic visuals and researching on different methods. Mapbox offers 50,000 free map loads and 200,000 free vector tiles API, which is more than enough for the scope of this project. Another advantage of using Mapbox is its versatility. It can be shared with numerous third parties including Tableau, ArcGIS Online, CARTO, and Fulcrum. This means this map can be used for further processing or be combined with other statistical data as one of the components of a dashboard (if using Tableau). For example, Statistics Canada's immigration statistics can be shown as line or bar graphs with projected population or percentage. It can also be developed using Android, Unity, and iOS.

Due to the sheer volume and size of the datasets – some as large as 35MB, data manipulation would not have been possible through Microsoft Excel or QGIS. In this case, RStudio was chosen to modify the CSV and shapefile datasets. The R script generated for this project can be simplified by using a function or loop during the process of calculating each province's attribute data in the CSV files. To promote reproducibility, the provincial recent immigrant percentages can be calculated and linked using RStudio as well (they were calculated separately and added manually in the current code). For simplicity and testing purposes, these aspects were not addressed in the project. It should be noted that python could have also been used to achieve the same results.

QGIS' utility in this project is to compress the heat map layer shapefile into a small enough size so that it can be uploaded into Mapbox while avoiding timeouts and layer visibility concerns at specific zoom levels. This process could have been achieved through other tools such as ArcGIS (subscription based but 1 year free for institutions with site licence) and mapshaper. The advantage of using QGIS is it is a free, open software with many enhanced capabilities. Without attempting to merge this project's CSV files with the CSD shapefile in QGIS, it would not have been known that this would result in a timeout response. Even though QGIS was not a definite requirement for the outcome of this project, it still has many great geographic information editing features which may become useful in future explorations. Therefore, ArcGIS or QGIS skills are definitely welcoming assets for future developers.

Heat Map Layer

One of the prior IRCC employee's comments is the map does not address gaps in supply and demand for service providers. From the government's perspective, service level gaps form the basis of funding allocation for different regions and locations. The estimated immigrant population map layer is this project's attempt to address this deficiency. Nevertheless, this data has not been validated sufficiently to draw meaningful conclusions regarding service gaps. Actual newcomer settlement locations may not align with Canadian's choice of settlement

location considering the differences in their backgrounds and interests. Information on newcomer settlement addresses need to be collected in order to validate this model's assumption and reveal actual service gaps.

The estimated immigrant population was first determined by FSA (Forward Sortation Area), which is the first three digits of the postal code. However, the outcome of this method did not appear logical. For instance, the map showed a strong concentration of newcomers in northern Alberta when in fact it was only the area around Fort McMurray that had the strongest population concentration. This meant the initial choice of using FSA to represent population concentration area was not ideal due to the large amount of surface area it covers. Thus, CSD was used instead to present area boundaries in more granularity. This method worked well because now the predicted population concentrations correspond much closer to reality and one's expectation. To improve this even more, a map layer could be created out of a shapefile that captures population data based on equally granular geographic surface areas (if this file exists). Furthermore, a feature that shows population statistics on hover of each CSD could be added as an enhancement.

Quebec's Uniqueness

According to the 2018 Annual Report to Parliament on Immigration, "Quebec has full responsibility for the selection of immigrants (except Family Class and in-Canada refugee claimants), as well as sole responsibility for delivering reception and integration services, supported by an annual grant from the federal government.)" under the Canada-Québec Accord. Thus, Quebec established its own 28 types of newcomer service provider categories. Mapping of Quebec's service types against that of federal's 13 types was a subjective process. In order to maintain a consistent and sustainable national database, it is recommended for Quebec to either adopt federal's 13 categories or create an objective, provincially derived mapping to these categories.

Conclusion and Future Work

To take this project a step further, a mobile application could be created. Mapbox is already mobile capable and with the ubiquity of mobile users, this will become a very useful feature. Another way this model can be improved is by integrating additional services which may be of interest to immigrants. Newcomers with children may be interested in finding top-notch education institutions for their children while those come with skill may be concerned with moving to a region with ample employment opportunities in their field of profession. Also, several users recommended housing services to become its standalone category because it is an essential need to which everyone will require access.

Different user profiles can be created for each stakeholder as well. The visual dashboard can be tailored to each type of stakeholder based on their own needs and interests. For example, statistics on immigration trend can be displayed for government users while only the service provider inventory and driving directions will be visible to the average newcomer.

One of the feedbacks received but was left unaddressed is displaying waitlists or real-time capacity levels of service providers. This is another way of demonstrating supply and demand equilibrium and can be explored by future works. It will require a new database which captures service provider capacity information before it can be visually presented. Furthermore, service providers will be required to constantly update the database in real-time. It would be ideal to allow users to directly register for services as well.

Finally, ratio of number of immigrant services to estimated number of immigrants per CSD (or other geographical boundary types) is another way to identify potential service level gaps (under the assumption that this model's calculation method is valid). A rough estimate of provincial ratio was calculated but it was not presented on the map. A new map layer could be created to visually represent the ratios while highlighting areas of concern (locations with low ratios). Even if a map layer is not created, the top five most concerning regions can be brought forth on a dashboard. This information will not only be of interest to government, but to service providers when they are deciding on locations to establish their services as well. To proceed even further, gap analysis could be done based on the assumption that a service location can provide services to those within 30 minutes of driving distance.

References

- Canada, G. o. (2019, April 8). *Call for proposals 2019: Settlement and Resettlement Assistance Programs funding guidelines*. Retrieved from Government of Canada: https://www.canada.ca/en/immigration-refugees-citizenship/corporate/partners-service-providers/funding/call-for-proposals-2019-settlement-and-resettlement-funding-guidelines.html#Eligible_services
- Canada, G. o. (2019, July 26). *Find free newcomer services near you*. Retrieved from Government of Canada: <http://www.cic.gc.ca/english/newcomers/services/index.asp>
- Canada, G. o. (2019, May 3). *Funding*. Retrieved from Government of Canada: <https://www.canada.ca/en/immigration-refugees-citizenship/corporate/partners-service-providers/funding.html>
- Canada, S. (2017, August 25). *Download, Census Profile, 2016 Census* . Retrieved from Statistics Canada: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/comp/page_dl-tc.cfm?Lang=E
- Canada, S. (2018, March 12). *2016 Census - Boundary files*. Retrieved from Statistics Canada: <https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm>
- Canada, S. (2019, February 20). *Immigration and Ethnocultural Diversity Highlight Tables*. Retrieved from Statistics Canada: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/imm/Table.cfm?Lang=E&T=11&Geo=00>
- Evans, S. a. (2017, Oct 10). *Mapped: How the US generates electricity*. Retrieved from CarbonBrief: <https://www.carbonbrief.org/mapped-how-the-us-generates-electricity>
- Immigration, D. e. (2018, March 1). *Carte des Partenaires*. Retrieved from Services offerts par les partenaires du Ministère: <https://services.immigration-quebec.gouv.qc.ca/fr/partenaires/carte-partenaires/>
- Immigration, R. a. (2018). *2018 ANNUAL REPORT to Parliament on Immigration*. Her Majesty the Queen in Right of Canada. Retrieved from <https://www.canada.ca/content/dam/ircc/migration/ircc/english/pdf/pub/annual-report-2018.pdf>

Appendix A

#	Federal	LINK		Quebec
A	Francophone service provider	A1-28	1	Accueil, installation, intégration, renseignement, référence
B	Help finding a job	B11	2	Activités de promotion
C	Job-specific language training		3	Aide technique
D	Help with daily life	D1, D4, D5, D6, D10, D19, D20, D21, D22, D23, D25, D26, D27, D28	4	Aide à la recherche de logement
E	Find or become a mentor to a newcomer	E14	5	Aide à la recherche de logement excluant les demandeurs d'asile
F	Services for refugees	D6, D11, F17	6	Aide à la recherche de logement incluant les demandeurs d'asile
G	Services for women		7	Concertation
H	Services for seniors		8	Francisation temps complet
I	Services for youth		9	Francisation temps partiel
J	Services for LGBTQ2		10	Information générale sur le logement
K	Language assessment		11	Information sur les possibilités d'emplois en région
L	Language training (general)	M8, M9	12	Installation et accompagnement des réfugiés pris en charge par l'État (RCPE)
M	Other services	M2, M3, M7, M12, M13, M15, M16, M17, M18, M24	13	Interprétariat et traduction
			14	Jumelage entre personnes immigrantes nouvellement arrivées et Québécoises et Québécois
			15	Mobilisation
			16	Médiation
			17	Préparation des candidats à l'établissement en région
			18	Rapprochement interculturel
			19	Sensibilisation
			20	Sensibilisation du milieu
			21	Service d'aide alimentaire (RCPE)
			22	Service d'aide vestimentaire (RCPE)
			23	Session Objectif intégration
			24	Sessions d'information (activités collectives)
			25	Soutien à l'accès aux services publics
			26	Suivi de l'intégration des candidats
			27	Suivi de l'intégration des candidats et des familles
			28	Séance Premières démarches d'installation