Generating measures of access to employment for Canada's eight largest urban regions

Technical Report (V1)

Corresponding Author
Jeff Allen
Department of Geography and Planning, University of Toronto St. George
100 St. George St., Toronto, Ontario M5S 3G3, Canada
jeff.allen@utoronto.ca

Steven Farber
Department of Human Geography, University of Toronto Scarborough
1265 Military Trail, Toronto, Ontario, M1C 1A4, Canada
steven.farber@utoronto.ca

Introduction

We create and release a publicly available dataset of neighbourhood level measures of access to employment for the eight largest urban regions in Canada. Measures of access to employment are key indicators for analyzing the characteristics of transport networks and urban form. Specifically, we generate cumulative measures (number of jobs reachable within 30, 45, and 60 minute commutes), gravity measures, as well as a competitive measure of accessibility which is standardized to allow for comparisons between regions. These are generated at the census Dissemination Area level for two travel modes, car and transit, including accounting for minute-by-minute variations in transit schedules. We release the data, and the code to generate it, openly on GitHub (`https://github.com/SAUSy-Lab/canada-transit-access`), as well as visualize the data on an interactive map (`https://sausy-lab.github.io/canada-transit-access/map.html`) so that they can easily be used by researchers, planners, and the general public. The input data and tools used are all open source so they can be shared or replicated elsewhere with minimal cost.

Input Data

Study Regions

We generate measures of access to employment for the eight largest urban regions in Canada: Toronto, Montreal, Vancouver, Calgary, Ottawa, Edmonton, Quebec City, and Winnipeg. We use Census Metropolitan Areas (CMA) for our study areas. CMAs are agglomerations of municipalities which pertain to urban areas with a population of 100,000 or more where at least 50% of the employed labour force works in the region's core, as determined from commuting data from the previous census (Statistics Canada, 2016a). Although not perfect, this measurement provides consistency of what constitutes the boundaries of urban regions across Canada. For our analysis, any adjacent CMAs are merged into one urban region due to the commuting flow and transit agencies that link adjacent regions. Two periphery CMAs within the Toronto region, Brantford and Peterborough, were not included as they did not have transit schedules available in a machine readable data format.

Demographic & Employment Data

For each of these eight regions, we use 2016 census Dissemination Areas (DA) to model the home locations of the labour force. DAs are the smallest areas in which socio-economic data is available from the quinquennial Canadian census, minimizing error due to the modifiable areal unit problem (see Kwan and Weber (2008) for a discussion of MAUP and its effects in accessibility research). DAs are designed and delineated for populations of 400 to 700 persons (Statistics Canada, 2016a), and have been used in other studies on transit accessibility in Canada (Widener et al., 2017; Wessel, Allen, & Farber, 2017). Specifically, we use the population weighted centroids of DAs snapped to the closest walking network segment to model the home locations of residents. Larger, neighbourhood sized Census Tracts (CT), however, are used for the location of employment, as they are the lowest level in which complete employment data was available for the 2016 census. It should be noted that several of these urban regions also run their own travel surveys (e.g. the Transportation Tomorrow Survey in the Toronto Region) with home and employment locations of residents, but we required data collected with consistent methodology across the country. Regional travel surveys typically have much more detailed travel diaries, but survey a lower percent of the overall population. The long-form census, which we draw our data from, is a 25% representative sample of Canadian households.

Network Graphs

Another primary input into our analysis are travel times linking where people live and places of employment. To compute these travel times, we built custom multi-modal network graphs for each urban region. Graphs for measuring transit travel times were built using the open-source routing engine OpenTripPlanner (2017). This has two sets of inputs. The first are the walking networks in each of these cities via the topological edges from OpenStreetMap. The second are transit schedules in the form of GTFS (Genreal Transit Feed Specification) data for every transit agency that serves these urban regions, circa May 2016 in order to align with the collection dates of the 2016 census. OTP uses the A^* algorithm to find shortest-path transit itineraries between each origin and destination. These graphs are inclusive of the time walking to and from stops, wait times, in-vehicle travels times, and transfers. It should be noted that GTFS represents the expected schedules generated by transit agencies, while the on-the-ground service of vehicles often differs from the schedule, and can potentially effect accessibility measures in some urban areas (Wessel et al., 2017). Real-time GPS data of transit vehicles is not available for all the agencies in our study regions, so this was not feasible for this project.

To provide comparison to transit travel times, we also compute travel times by driving, using OpenStreetMap data as the input network. The travel times for driving were computed with a different routing engine, Open Source Routing Machine (OSRM) (Luxen & Vetter, 2011), as it includes greater consideration for driving attributes like speed limits, turn restrictions, and one-way streets.

Table 1: Summary of CMAs and Transit Agencies used in this study

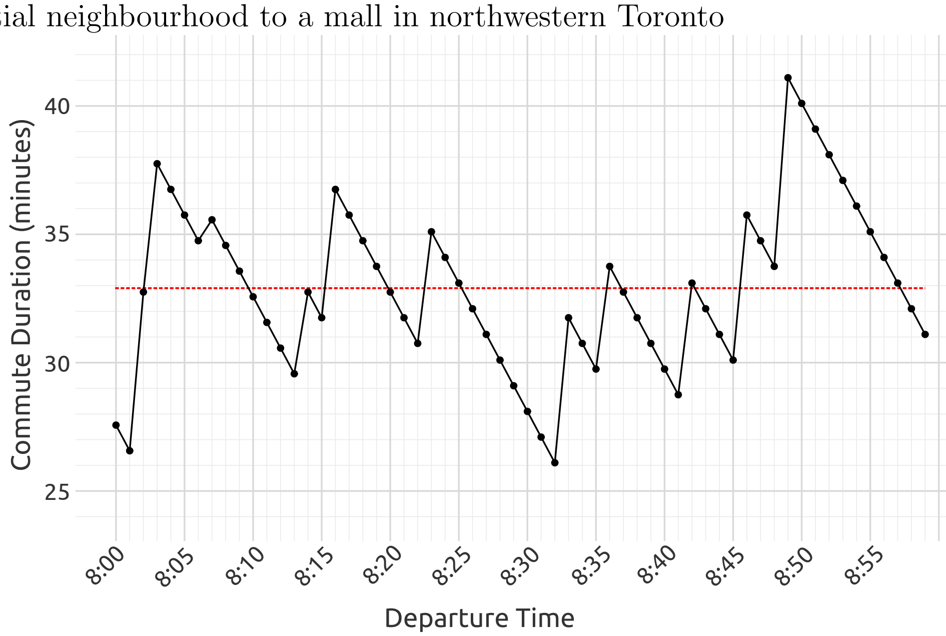
	Area (km^2)	Population	CMAs	Transit Agencies
Toronto	12,160	7,951,192	Oshawa, Toronto, Hamilton, St. Catharines-Niagara, Kitchener-Cambridge-Waterloo, Guelph, Barrie	Toronto Transit Commission, Durham Regional Transit, GO Transit, York/VIVA, MiWay, Brampton Transit, Oakville Transit, Burlington Transit, Hamilton Street Railway, Niagara Region Transit, Guelph Transit, Barrie Transit, Grand River Transit, Toronto Island Ferries
Montreal	4,605	4,098,927	Montreal	Agence mtropolitaine de transport, CIT Chambly-Richelieu-Carigna, CIT du Haut-Saint-Laurent, CIT La Presqu'le, CIT Laurentides, CIT Le Richelain, CIT Roussillon, CIT Sorel-Varennes, CIT Valle-du-Richelieu, CRT Lanaudire, MRC de Deux-Montagnes, MRC de L'Assomption, MRC Les Moulins (Urbis), Rseau de transport de Longueuil, RTM Sud-ouest, Societe de transport de Laval, Socit de transport de Montral, OMIT Sainte-Julie
Vancouver	4,935	2,745,461	Vancouver, Abbotsford-Mission, Chilliwak	BC Transit, TransLink, West Coast Express
Calgary	5,110	1,392,609	Calgary	Calgary Transit, Airdie Transit
Ottawa	6,770	1,323,783	Ottawa - Gatineau	OC Transpo, Socit de transport de l'Outaouais
Edmonton	9,440	1,321,426	Edmonton	Edmonton Transit Service, Fort Sask Transit, St. Albert Transit, Strathcona County Transi
Quebec City	3,410	800,296	Quebec City	Rseau de transport de la Capitale, Socit de transport de Lvis
Winnipeg	4,310	778,489	Winnipeg	Winnipeg Transit

Measuring Access to Employment

Computing Travel Times

The first step of our analysis was to compute travel time matrices of DAs (home locations) to CTs (employment locations) for each of the eight urban regions in our study. Because of the inherent temporal variations in transit schedules, we follow the precedent in the literature to compute transit travel times for every minute of the morning commute period (Owen & Levinson, 2015; Farber & Fu, 2017), to be subsequently averaged when computing accessibility metrics. Figure 1 exemplifies how travel time by transit between a residential neighbourhood and an employment centre can vary substantially, and selecting one of these travel times could greatly over- or under-estimate the travel time during this period.

Figure 1: Example of the temporal differences in commute time by transit from a residential neighbourhood to a mall in northwestern Toronto



For our analysis, this was computed in parallel over several processing units which output results for multiple departure times, τ. The outputs are stored in a three-dimensional array.

$$T_{i,j,\tau} = \left\{t_{i,j,\tau}\right\} \tag{1}$$

Where each cell, $t_{i,j,\tau}$, is the travel from the origin DA, i, to the destination CT, j, for a specific departure time, τ. Due to heavy computation, travel times were capped at 90 minutes, assuming that no one would be willing to travel to jobs that require more than a 90 minute commute.

Due to a lack of openly available network level congestion data, travel times for driving were computed as free-flow speeds, and then multiplied by a congestion factor, k_c, to account for how peak-hour travel is slower than off-peak. The congestion factors were set at 1.7 for Toronto and Vancouver, 1.6 for Montreal, 1.5 for Ottawa, and 1.4 for the remaining four cities. These values were estimated from reports examining costs of congestion in Canadian cities (Metrolinx, 2008; Urban Transportation Task Force, 2012) as well as from the private data vendor TomTom, which hosts an online worldwide ranking of congestion by city (TomTom, 2018). We also apply a minor two minute penalty for parking, t_p. The peak hour travel time by driving between two locations, $t^*_{i,j,d}$, is thus calculated from the free flow travel time, $t_{i,j,d}$, as follows.

$$t^*_{i,j,d} = k_c \, t_{i,j,d} + t_p \tag{2}$$

Cumulative Access Measures

The first, and simplest, measure of accessibility we compute is cumulative accessibility. This is the count of employment opportunities that can be reached within a specified travel time, θ, and is formulated as follows

$$A_{i,\theta} = \sum_{j=1}^{J} O_j f(t_{i,j}, \theta) \tag{3}$$

O_j is the number of job opportunities at location j. $f(t_{i,j}, \theta)$ is a binary function of whether the travel time from i to j is less than a travel time threshold, θ.

$$f(t_{i,j}, \theta) = \begin{cases} 1 & \text{if} \quad t_{i,j} \leq \theta \\ 0 & \text{if} \quad t_{i,j} > \theta \end{cases} \tag{4}$$

Because of the inherent continuous temporal variations in transit schedules, we average these measures over the morning rush hour period (from τ_a to τ_b)

$$\bar{A}_{i,\theta} = |\tau_b - \tau_a|^{-1} \int_{\tau_a}^{\tau_b} A_{i,\theta}(\tau) d\tau \tag{5}$$

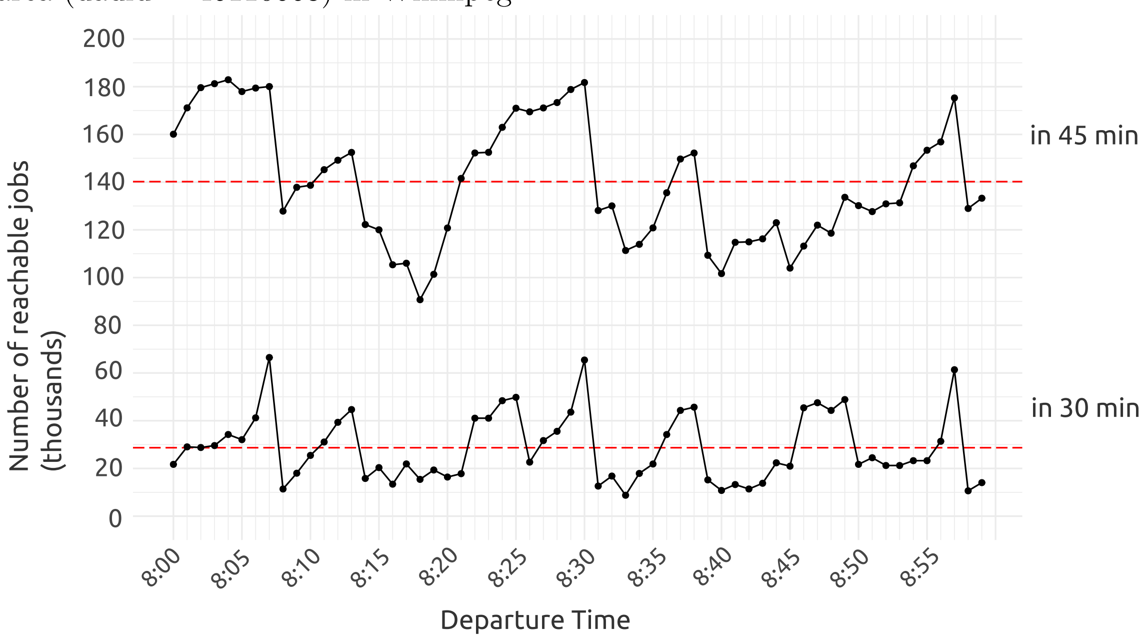
Since we computed travel times at a per minute basis, $\bar{A}_{i,\theta}$ can be generalized as follows.

$$\bar{A}_{i,\theta} = |120|^{-1} \sum_{\tau \in M} \sum_{j=1}^{J} O_j f(t_{i,j,\tau,\theta}) \tag{6}$$

Where M is every minute τ from 7:00am to 8:59am. Figure 2 exemplifies how departure time can seriously effect accessibility measures, and why averaging is beneficial.

The output can also be highly sensitive to θ, particularly for $t_{i,j}$ which are close to the threshold. For example, for $\theta = 30$, an opportunity 29 minutes away is counted, but an opportunity 31 minutes away is not, even though the difference between these is only two minutes travel time. We computed cumulative accessibility for $\theta = 30$ minutes, $\theta = 45$ minutes, and $\theta = 60$ minutes to allow for comparison.

Figure 2: Differences in cumulative accessibility by departure time for a dissemination area (dauid = 46110663) in Winnipeg



Gravity Access Measures

While cumulative accessibility measures are relatively simple to understand, they do not account for how job opportunities nearby are more attractive than those further away due to the time savings resulting from reduced commute durations. To account for this, we also compute gravity measures of accessibility where the function, $f(t_{i,j})$, weights nearby opportunities more than those that are further away via a decay function. For our study, we compute access to jobs using an inverse-power function, parametrized such that a 30 minute commute returns a value of 0.5, and with a maximum value of 1 (at $t_{i,j} = 0$). 30 minutes is approximately the average commute duration across all eight regions (Statistics Canada, 2016b).

$$A_i = \sum_{j=1}^{J} O_j f(t_{i,j}) \tag{7}$$

$$f(t_{i,j}) = 180(90 + t_{i,j})^{-1} - 1 \tag{8}$$

Competitive Access Measures

Gravity and cumulative measures of access to jobs are inadequate when comparing results between cities because they do not account for the size and spatial distribution of the labour market which competes for employment opportunities (Shen, 1998; Geurs & van Eck, 2003).

To account for this, we also computed competitive measures of accessibility. This technique accounts for access at both the demand and supply locations of analysis (Weibull, 1976), and has been commonly used in access to health services (Luo & Wang, 2003; Delamater, 2013). Applied to access to employment, this accounts for how employment opportunities and the labour force are both spatially distributed and overlapping, and that competition exists among the labour force for employment opportunities (Shen, 1998; Geurs & van Eck, 2003). Moreover, competitive accessibility measures have been shown to be a better predictor of employment outcomes than accessibility measures that do not consider competition (Merlin & Hu, 2017). Mathematically, this technique involves normalizing employment opportunities at j by their labour market catchment area, L_j (i.e. this is the demand for jobs at j), and solving iteratively where P_i is the size of the labour force at i.

$$A_i = \sum_{j=1}^{J} \frac{O_j f(t_{i,j})}{L_j} \tag{9}$$

$$L_j = \sum_{i=1}^{I} \frac{P_i f(t_{i,j})}{A_i} \tag{10}$$

This process also accounts for how each P_i has varying levels of access, but can only fill a set amount of jobs. i.e. employers compete for workers who have varying levels of access to jobs, just as people compete for jobs at locations which have varying access to the labour force (Geurs & van Eck, 2003; Merlin & Hu, 2017).

For our study of Canadian cities, we expand the above formulas to account for a labour force which commutes by car or by transit, including averaging transit over the morning commute period (7:00am to 8:59am) because of fluctuations in the transit schedules.

$$A_{i,T} = k|120|^{-1} \sum_{\tau \in M} \sum_{j=1}^{J} \frac{O_j f(t_{i,j,\tau}) f(t_{i,j,\tau})}{L_j} \tag{11}$$

$$A_{i,D} = k \sum_{j=1}^{J} \frac{O_j f(t_{i,j,d}) f(t_{i,j,d})}{L_j} \tag{12}$$

$$L_j = |120|^{-1} \sum_{\tau \in M} \sum_{i=1}^{I} \frac{\alpha_{i,T} P_i f(t_{i,j,\tau})}{A_{T,i}} + \sum_{i=1}^{I} \frac{\alpha_{i,D} P_i f(t_{i,j,d})}{A_{D,i}} \tag{13}$$

$A_{i,T}$ is the accessibility measure for transit, and $A_{i,D}$ for driving. $t_{i,j,d}$ is the travel time by driving during the commute period. The impedance functions for transit and driving, $f(t_{i,j,\tau})$ and $f(t_{i,j,d})$, use the inverse-power function presented in (8). k is a scaling factor. $\alpha_{i,D}$ is the commute mode share ratio of workers at location i who travel to work via private vehicle. $\alpha_{T,i}$ is the mode share ratio by transit and walking. The mode share for transit for our study is assumed as the total non-driving commuting population ($\alpha_{i,T} = 1$ - $\alpha_{i,D}$), and therefore also includes the small percent of those who take active modes (bike or walk). This assumes that those who bike or walk to work are also able to commute to work by transit, but not by car. The two $f(t)$ terms in A_i permits accurate comparisons between cities which have differing transport networks and sub-optimal distribution of opportunities (Delamater, 2013). The resulting values of $A_{i,T}$ and $A_{i,D}$ are scaled (via the parameter k) from 0 to 1 to provide easier interpretation, where 0 is no access and 1 is the maximum level of access to employment observed for any travel mode across Canada.

For thousands of zones, and minute-by-minute travel times, the process for computing multiple iterations of competitive accessibility is computationally intensive. We therefore examined convergence at a regional scale for Winnipeg to see what a suitable point would be to exit the loop. With only three iterations, the average absolute difference in A_i is less than 0.1% compared to the previous iteration, which we deem suitable to stop and refrain from future iterations.

Disseminating & Visualizing Results

The output data as well as the code used to compute travel time matrices and different accessibility measures are be publicly available on Github (`https://github.com/SAUSy-Lab/canada-transit-access`).

Since the computed accessibility measures are linked to areal units, they can be visualized as choropleths to examine their spatial patterns. We built an interactive map to display and share our accessibility measures (`https://sausy-lab.github.io/canada-transit-access/map.html`). This map allows for switching between cities, selecting and comparing access by travel mode, and comparing between the different types of access measures generated (cumulative, gravity, and competitive). The map also includes the option of overlaying the locations of different demographic groups as a dot density layer to examine how and where different groups are aligned with regions of low access.

Descriptive Results

Comparative results for the measure of competitive accessibility scaled from 0 (lowest) to 1 (highest) are displayed in Table 2. The maximum values in Table 2 provide a sense of how the best served areas in cities compare with each other. We tabulate data for both transit access and auto access, as well as a ratio between transit and auto access, to examine the differences between these two modes. Figures 3 and 4 are plots of the distribution of access to examine how clustered or dispersed values are from the mean for each region.

Table 2: Summary statistics of access to jobs by mode and urban region

	$A_{i,T}$		$A_{i,C}$		$A_{i,T}/A_{i,C}$	
	Mean	Max	Mean	Max	Mean	Max
Toronto	0.094	0.609	0.378	0.994	0.214	0.722
Montreal	0.097	0.466	0.422	0.912	0.189	0.598
Vancouver	0.135	0.625	0.384	0.848	0.285	0.752
Calgary	0.081	0.373	0.404	0.782	0.174	0.501
Ottawa	0.119	0.480	0.518	1.000	0.201	0.483
Edmonton	0.070	0.337	0.402	0.705	0.149	0.489
Quebec City	0.104	0.329	0.537	0.829	0.172	0.429
Winnipeg	0.133	0.387	0.540	0.800	0.230	0.516
All	0.101	0.625	0.411	1.000	0.210	0.752

Figure 3: Plot indicating the mean and distribution of access to jobs by transit

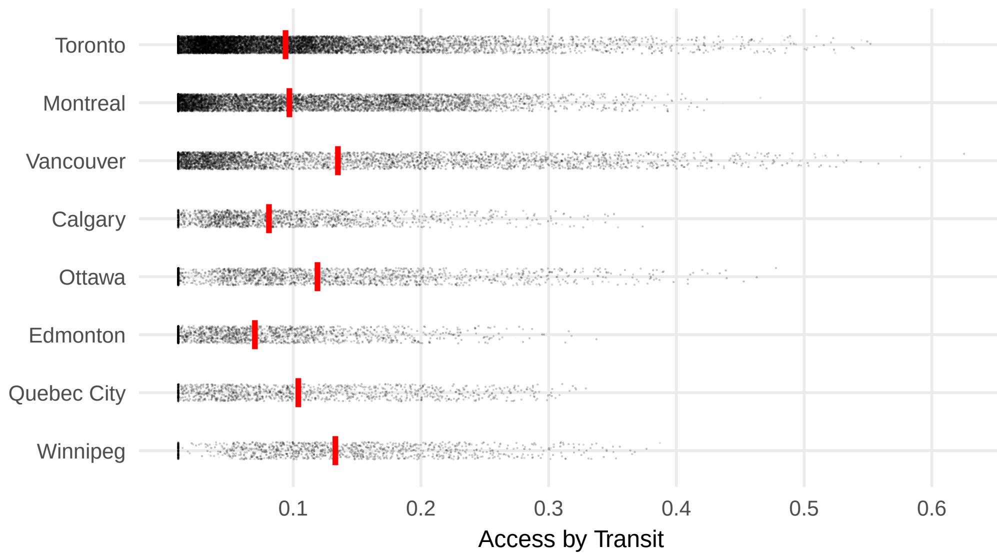
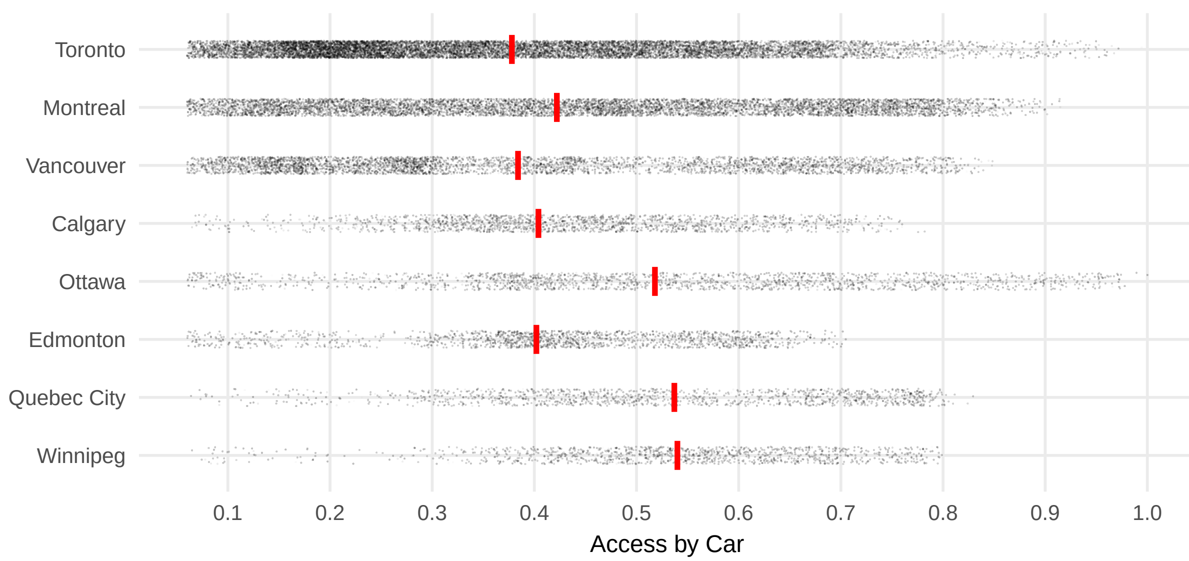


Figure 4: Plot indicating the mean and distribution of access to jobs by car



References

Delamater, P. L. (2013). Spatial accessibility in suboptimally configured health care systems: a modified two-step floating catchment area (m2sfca) metric. *Health & place*, *24*, 30–43. doi: 10.1016/j.healthplace.2013.07.012

Farber, S., & Fu, L. (2017). Dynamic public transit accessibility using travel time cubes: Comparing the effects of infrastructure (dis) investments over time. *Computers, Environment and Urban Systems*, *62*, 30–40. doi: 10.1016/j.compenvurbsys.2016.10.005

Geurs, K. T., & van Eck, J. R. R. (2003). Evaluation of accessibility impacts of land-use scenarios: the implications of job competition, land-use, and infrastructure developments for the netherlands. *Environment and Planning B: Planning and Design*, *30*(1), 69–87. doi: 10.1068/b12940

Kwan, M.-P., & Weber, J. (2008). Scale and accessibility: Implications for the analysis of land use–travel interaction. *Applied Geography*, *28*(2), 110–123. doi: 10.1016/j.apgeog.2007.07.002

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a gis environment: synthesis and a case study in the chicago region. *Environment and Planning B: Planning and Design*, *30*(6), 865–884. doi: 10.1068/b29120

Luxen, D., & Vetter, C. (2011). Real-time routing with OpenStreetMap data. In *Proceedings of the 19th acm sigspatial international conference on advances in geographic information systems* (pp. 513–516). New York, NY, USA: ACM. doi: 10.1145/2093973.2094062

Merlin, L. A., & Hu, L. (2017). Does competition matter in measures of job accessibility? explaining employment in los angeles. *Journal of Transport Geography*, *64*, 77–88. doi: 10.1016/j.jtrangeo.2017.08.009

Metrolinx. (2008). *Costs of road congestion in the greater toronto and hamilton area: Impact and cost benefit analysis of the metrolinx draft regional transportation plan* (Tech. Rep.).

OpenTripPlanner. (2017). (http://www.opentripplanner.org/)

Owen, A., & Levinson, D. M. (2015). Modeling the commute mode share of transit using continuous accessibility to jobs. *Transportation Research Part A: Policy and Practice*, *74*, 110–122. doi: 10.1016/j.tra.2015.02.002

Shen, Q. (1998). Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers. *Environment and planning B: Planning and Design*, *25*(3), 345–365. doi: 10.1068/b250345

Statistics Canada. (2016a). *Census dictionary.*

Statistics Canada. (2016b). Census of population.

TomTom. (2018). *TomTom Traffic Index.* Retrieved from `https://www.tomtom.com/en_gb/trafficindex/`

Urban Transportation Task Force. (2012). *The high cost of congestion in canadian cities* (Tech. Rep.). Council of Ministers Responsible for Transportation and

Highway Safety.
Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. *Regional science and urban economics*, *6*(4), 357–379. doi: 10.1016/0166-0462(76)90031-4
Wessel, N., Allen, J., & Farber, S. (2017). Constructing a routable retrospective transit timetable from a real-time vehicle location feed and gtfs. *Journal of Transport Geography*, *62*, 92–97. doi: 10.1016/j.jtrangeo.2017.04.012
Widener, M. J., Minaker, L., Farber, S., Allen, J., Vitali, B., Coleman, P. C., & Cook, B. (2017). How do changes in the daily food and transportation environments affect grocery store accessibility? *Applied Geography*, *83*, 46–62. doi: 10.1016/j.apgeog.2017.03.018