A High Performing, Scalable Model for Computing and Visualizing Public Transit Accessibility

A Case Study on Cultural and Art Amenities in Metro Vancouver

2021 Capstone Project

University of British Columbia, Okanagan

Luka Vukovic

Rain Shen

Yuxuan Cui

Graham Kerford

THE UNIVERSITY OF BRITISH COLUMBIA

Statistics Canada

Statistique Canada

Table of Contents

List of Abbreviations

GTFS	General Transit Feed Specification
OSM	Open Street Map
R5	Rapid Realistic Routing on Real-World Reimagined networks

List of Tables

List of Figures

Executive Summary

Dashboard Link: Vancouver Transit Accessibility to Cultural Amenities Dashboard

Introduction

Transportation network analysis is fundamental to urban planning for it determines how resources are distributed across a population. Resources come in the form of amenities such as grocery stores, schools, parks, and hospitals. Our client, Statistics Canada produces data to better understand Canada's population, resources, economy, society, and culture. They have previously developed network accessibility measures based on distance of driving, and walking to compute proximity scores for various types of amenities.

Problem

Accessibility measures based on time using transit have not yet been incorporated into proximity scores due to its multi-modal complexity and computational intensity. In 2016, 22.3% of Canadians depended on public transit in large cities; thus, incorporating transit accessibility measures is paramount to not under-represent large segments of the population which can inevitably worsen pre-existing inequalities in the urban landscape.

Objective

The aim of this project was to establish a first iteration of an open source scalable framework for data collection and analysis of transit accessibility measures. We validated our framework on Vancouver, raising the question of, "How accessible are Vancouver's cultural amenities (libraries, museums, art galleries, and theatres) using the current transit system?"

Methodology/Results

To address the computational intensity of multimodal shortest path routing, we use Conveyal's R5 realistic routing algorithm available in R as r5r. It allows us to compute over 5.3 million transit routes repeatedly, 360 times in a day over 3 days, in just a matter of one hour. The travel time matrix was then used to develop three accessibility measures: one based on time, one on scores, and one on percentiles which were visualized with Leaflet and Kepler.gl and embedded in an R shiny dashboard.

Conclusion

This project provides a high performing and scalable framework for producing three unique transit accessibility measures for network analysis using Greater Vancouver as an initial use-case scenario. The frameworks can be further developed and adopted by urban developers to ensure equitable, sustainable, and optimal urban design for years to come.

1 Introduction

Transportation network analysis is crucial to urban planners, researchers, and policy makers to ensure accessibility remains a key feature of the urban landscape. This requires both a realistic and a scalable measure for accessibility, which can impact the planning and development of roads and city districts, while also helping analyze a population's access to healthcare, schools, grocery stores, and other amenities. On a fundamental level, having proper accessibility measures allows urban resource distribution to be more equitable, optimal, and based on informed decisions.

Despite the importance of such measures, much of past accessibility research only considers a limited set of travel modes. These include driving, biking, and walking, but overlook the importance of public transit as a primary mode of travel (Liu and Zhu, 2004). This may be for a few reasons such as a lack of standardized public transit data, or simply a lack of routing engines that can efficiently model complex journeys through transit networks.

Measuring transit accessibility is crucial and will be more than ever for a few major reasons. First, being in the midst of a *Third Industrial Revolution*, much focus will be diverted to urban and regional planning (Roberts, 2015). Sharing economies and sustainability driven planning will undoubtedly increase our dependence on public transit and electric vehicles, although it is uncertain what fraction of society will own such technology within the next few decades. Instead, one might find access to electric vehicles through autonomous ridesharing systems, which will more than likely be useful for filling holes in current public transit networks (Mo *et al.*, 2021). Not to mention, car dependence generally has poor outcomes on physical health, psychological health, and the environment through traffic congestion, substantial loss of time, physical inactivity, and accidents (Martin *et al.*, 2014; Royal Society for Public Health, 2016; Sallis *et al.*, 2004; EMBARQ, 2013). As cities grow and our interconnectivity increases, the importance of having a transit network benchmarking framework remains vital for supporting urban growth.

To better understand the scale of public transportation in Canada, prior to the COVID-19 pandemic, 31.4% of Canadians regularly used sustainable transportation, where public transit comprised almost 40% of those cases. When considering Canada's largest three metropolitan areas, as much as 40.4% of the population used sustainable transportation, of which 55% was public transit (Statistics Canada Census, 2016). To say the least, public transit will remain an imperative area for urban developers to focus on.

Secondly, we must recognize that certain population segments do not readily have access to private transportation. These populations also tend to be more vulnerable, whether they be marginalized, elderly, or youth, and are likely to have greater reliance on public transit relative to other social segments. Further, inequitable distribution of transit accessibility only exacerbates pre existing inequalities (Lubitow *et al.*, 2017).

As such, measuring, visualizing, and analyzing transit accessibility across cities is paramount for equitable urban planning and growth.

To address the lack of standardized methods for obtaining transit accessibility information, this project responds with a first iteration methodology for a simple, scalable, and high performing framework for data generation, analysis, and visualization. Given a road network, a transit network, and coordinates of every city block for a city the size of Vancouver, within a matter of 1 hour over 5.7 billion computations could be performed to generate a many-to-many point travel time matrix which can then be manipulated for network analysis. The greater metropolitan area of Vancouver was used as a proof of concept example in the context of measuring transit access to cultural amenities such as museums, libraries, art galleries, and theatres.

2 Background

2.1 Measurement of Accessibility

The concept of accessibility was first brought up by Hansen (1959) as “the potential of opportunities for interaction”. As such, accessibility refers to the physical access to goods, services, activities, and destinations (Saif *et al.*, 2019). It provides a measurement of the number of the destinations that can be reached from a specific point (Deboosere and El-Geneidy, 2018). In the context of urban systems, accessibility is an essential component of the transportation network. In general, there are three common metrics of accessibility. First one is measuring the number of destinations which can be reached from a certain point within a time threshold using one specific transportation mode (Morris *et al*., 1979). Second one is measuring the destination by distance, in which the further the destination is, the less accessible it is (Koenig, 1980). The last one is assigning a certain utility to each destination and computing the logsum of all destinations (Handy and Niemeier, 1997). Among these three metrics, the first one is most interpretable and easiest to compute (Deboosere and El-Geneidy, 2018). According to the metrics, in the most recent studies, accessibility is usually measured by the distance between the origin and the destination, or by the time the trip takes from the origin to the destination (Cheng and Chen 2015, Saghapour *et al*., 2016).

2.2 Public Transit

In the past few decades, the way of transportation has gradually moved from private vehicles to more sustainable modes, for example, public transit, walking, and cycling (Albacete *et al*., 2017). In Canada, the percentage of the population who use public transit as the primary transportation mode has significantly increased since 1996, especially in the large metropolitan areas (Statistics Canada, 2017). In the largest metropolitan cities, like Toronto, Montreal, and Vancouver, approximately 22.3% of the population use public transit for their daily commuting (Statistics Canada, 2017). A well-developed public transit system not only increases urban connectivity to resources, but also improves social inclusion (Said *et al*., 2019). Therefore, being able to estimate transportation efficiency can allow urban planners to directly act on maximizing social opportunity within the urban economic system (Wang and Chen, 2015).

2.3 Cultural and Art Amenities

Cultural and art amenities are a particular type of amenity that increases quality of life in communities by making them more attractive places to live (Sheppard, 2014). Hence, measuring accessibility to cultural and art amenities is an effective way to assess social opportunity and livelihood across local communities. There are numerous studies focusing on this task in the context of other amenity types. Particularly, in 2007 Burns and Inglis calculated travel time to supermarkets and fast-food outlets by driving, walking, and public transit. Travel time on public transit was computed according to the frequency of buses and the types of roads, which is not the most realistic approach since it does not directly measure transit time through a network. Generally however, there have been limited studies regarding cultural and art amenities, hence why they were chosen for a proof of concept accessibility framework.

Unique to cultural amenities, their '*mass*' becomes an essential component to consider when modeling accessibility. In a study from Statistics Canada, *mass* can be thought as the amount of service an amenity provides (Statistics Canada, 2021). Regarding cultural amenities, for example, a library with larger space and higher volume provides services to more people which should be weighted more than smaller libraries. Similarly, popular public galleries would be expected to hold more visitors than private ones. In this project, we provide a new index to account for the *mass* of amenities and finally describe its impact on accessibility via public transit.

3 Methodology

In this section, we detail how the research problem was approached computationally. We explain what data a scalable model would require, and how transit accessibility scores are computed and visualized from the data in an efficient manner.

3.1 Framework Data

Data sources for computing and visualizing transit accessibility measures are summarized in Table 1.

Data	*Features*	*Source*
General Transit Feed Specification (GTFS)	All transit network data (stop coordinates, stop times, bus routes, etc.)	TransLink Open API[1]
Dissemination Blocks (Origins)	Unique block ID, latitude, and longitude of the city block's centroid	Census of Population[2]
Amenities (Destinations)	Unique destination ID, types, latitude, and longitude	The Open Database of Cultural and Art Facilities (ODCAF)[3]
OpenStreetMap (OSM)	Urban street network data	openstreetmap.org
Geospatial Shapefile	Dissemination block unique ID, longitude/latitude polygon data	Census Cartographic Boundary File[1]
'*Mass'*[1] and Importance Features on Amenities	Operating Hours/Days, Visitors, Ratings, Reviews; Total space, total volume	Google API and BC Public Libraries Systems

Table 1. Data required for many-to-many point travel time matrix computation and accessibility measure visualization. [1] *Mass* corresponds to the relative amount of services an amenity provides. [1] TransLink Open API (May 2021). [2] Statistics Canada (2016). [3] Statistics Canada (2020).

3.2 Computing Travel Time Matrices

To evaluate transit accessibility across an urban landscape, many-to-many point travel times need to be computed into a travel time matrix. This requires a street network (Open Street Maps), a transit network (GTFS), origin-destination coordinates, and a routing engine. Popular open-source routing engines include Open Trip Planner 1 (OTP-1), Open Trip Planner 2 (OTP-2), Conveyal's R5 (Rapid Realistic Routing on Real-world and Reimagined networks), and GraphHopper.

GraphHopper offers fast compute with many-to-many point matrices; however, it was not considered for the lack of examples on matrices larger than 100x100. Our matrix contained roughly 15,000 x 350 points.

OTP-1 and OTP-2 are generally focused on passenger facing journey planning. OTP-1 has analysis functionality but performs slow relative to other engines, using a generalized cost A* algorithm. OTP-2 is better optimized with a Multi-criteria range-RAPTOR algorithm but does not support one-to-many point routing analysis which is a necessity of this project. Most importantly however, both OTPs optimize routing on generalized cost instead of on minimizing travel time. For example, OTP may opt for a single longer bus ride than a shorter route with 2 or 3 transfers. In reality, transit users typically aim to minimize time, not cost.

R5 supports and is optimized for time-window trip planning which better reflects how people use the transportation system. R5, being implemented in Java, was also intended for analysis applications being magnitudes faster and less memory intensive than engines that are similar to OTP-1, particularly for one-to-many point routing and travel time matrix generation. (OTP website) As such, travel times were computed using the open source r5r library, an R implementation of Conveyal's R5 routing engine, with multimodal transit networks built from the GTFS and OpenStreetMap data detailed in 3.1. (Pereira, 2021) The final travel time matrix was an aggregation of 36 trips, departing every hour from 7:00am to 7:00pm with a 30-minute departure window on a weekday, a Saturday, and a Sunday (12 x 3). This allowed us to average travel times across changing bus schedules throughout the week.

3.3 Measuring Transit Accessibility

Transit accessibility to cultural amenities was measured in three fashions. The first was a direct time measure to create an isochrone map. These can be interpreted as equal-time (*iso-chrone*) contours surrounding the nearest amenity of interest, but since there are many amenities the there are essentially dozens, if not hundreds, of small isochrones simultaneously visualized on the same map. The second measure involved scoring each dissemination block's transit access using the worst case trip time to the nearest *n* amenities where *n* is how many amenities to consider (nearest 1, 2, 3, …, *n*) and worst case time is the inverse of the sum of mean transit time and two standard deviations (equation 1). If more than $n = 1$ amenities are considered, the score for each amenity is summed to yield an overall score (equation 2). When scores are summed they are normalized (optionally log normalized to normalize the distribution) to yield a final range of 0 - 1. The best (shortest) transit times yield scores closer to one, while the longest transit times yield scores closer to zero. The third measure was the percentile of these scores which uniformly distributes scores allowing for more interpretability than raw scores alone.

$$s_{o_i,d_j} = \frac{1}{\mu_{o_i,d_j} + 2\sigma_{o_i,d_j}}$$

Equation 1

$$s_{o_i} = \sum_{j=1}^{\{1,2,3,n\}} s_{o_i,d_j}$$

Equation 2

3.4 Weighing Amenity Importance (*Mass*)

The amenity weight index, or *mass*, reflects amenity importance. This is the amount of services an amenity provides, its popularity, and even physical capacity. Google API was leveraged using names and geolocation points from the ODCAF dataset to gather total reviews and rating data. The index for galleries, theatres, and museums is described by equation 3 which takes the mean of four normalized *mass* indicators: weekly hours, days open, ratings, total google review. For libraries (equation 4) three additional features, volume, books, and annual visitors were incorporated from the *BC Public Libraries Systems*. A higher weight index indicates the amenity has more importance. Feature normalization for equal weighing should be modified based on the use case and assumptions. .

$$weight_i = \frac{n(hours_i) + n(days_i) + n(rating_i) + n(reviews_i)}{N}$$

Equation 3

$$weight_i = \frac{n(hours_i) + n(days_i) + n(rating_i) + n(reviews_i) + n(volume_i) + n(books_i) + n(annual\ visitors_i)}{N}$$

Equation 4

3.5 Measuring Transit Efficiency

The efficiency score was calculated by taking the difference between accessibility score percentiles and the needs score percentile for each *i*th block (equation 5). The needs score is the mean of traffic intensity, population, and the amenity densities, all of which are normalized. The traffic intensity is based on traffic surveys taken by the British Columbia Government between 1997 and 2015 (Government of British Columbia, 2015). These surveys measure the average number of vehicles that pass an 8 min period measured every quarter of an hour. Since not every survey site is used each year, only data from the most recent year was used in the analysis. Due to the sparsity of the surveys, traffic intensity for blocks were derived by taking the mean vehicle count within a 5 km radius of each block and normalizing the data. The amenity density is the normalized value of all amenities within a 1 km walking distance of the center of each block assuming an average walking speed of 3.6 km/h.

$$Eff_i = \%(score_i) - \%(\mathrm{mean}(population_i,\ \ traffic_i,\ \ amenity\ density_i))$$

Equation 5

3.6 Visualization

Three primary tools were used for visualization, Pandana (Python), Kepler.gl (Python), and Leaflet (R). Pandana was used for visualizing the pedestrian and transit network by combining GTFS and OSM data to build a network model. Kepler.gl and Leaflet were used for choropleth visualizations of accessibility measures by layering city block shape files (which contained accessibility scores) onto Vancouver's OSM. Kepler allows for data filtering but this requires raw data to be built into the html render, which is not ideal for performance. Leaflet allows for 2D geospatial visualization and does not depend on data, allowing for lighter map files which could be rendered on a dashboard within a second, compared to slow performing live rendering. Choropleth maps were deployed on an R Shiny dashboard which was preferred over others due to its sleekness and ease of use.

4 Results and Discussions

4.1 Accessibility Measures

As detailed in Methodology, three types of transit accessibility measures were used: time based measures, scoring based measures, and score percentiles measures. Each measure offers a unique set of advantages, disadvantages, and unique means to analyze transit accessibility across an urban landscape summarized in table 2.

Advantages	Disadvantages
Score Measures - Easily incorporates the uncertainty in transit travel time. - Easily incorporates features beyond travel time such as transfers, modes of transportation, amenity popularity weights, etc. - Offers a standardized measure which can be useful for comparisons across different cities.	*Score Measures* - Lacks interpretability. - Can be highly skewed.
Percentile Measures - Offers high interpretability for comparing specific areas to the rest. - Data distribution is uniform. - Easily incorporates multiple scoring features and uncertainty measures.	*Percentile Measures* - Provides no information on the actual transit accessibility. - Information is lost on the true differences in accessibility between city blocks.
Time Measures - Absolute transit time offers the highest interpretability for general analysis of any city block.	*Time Measures* - Does not incorporate other features of accessibility. - Does not consider the uncertainty in transit time. - Not necessarily standardized for comparisons between cities.

Table 2. Advantages and disadvantages of the Primary Transit Network Accessibility Measures

The overall distribution of Vancouver's accessibility measures are summarized in Figure 1. We notice that in **1A** that the scores are highly positively skewed. This can be attributed to the inverse property and city blocks which either neighbour or include amenities. These have travel times of less than 1 minute. To avoid zero division in the formula (see section 3.3), these values were converted to 1 minute, which then gives a highest possible score of 1. Almost all other blocks were computed with travel times exceeding 10 minutes (from **1C**). As a result, most scores were below 0.1 as shown by equation 6 below.

$$t \leq 10,\ \sigma_t = 0 \qquad score \leq \frac{1}{t + 2\sigma_t} \leq 0.1$$

Equation 6

Despite score measures in **1A** providing a useful metric for standardizing accessibility to a range of 0 - 1, scores are not very interpretable, particularly when they're positively skewed. For interpretability, percentiles (**1B**) can be used. They render the original distribution uniform allowing for relative comparisons between blocks; however, they no longer provide information on true accessibility. For example, as an extreme case, a 95th percentile block may have better transit access than 95% of other city blocks yet the transit time to a specific amenity may still be very poor. This type of information loss is even less desirable if the original distribution is multimodal.

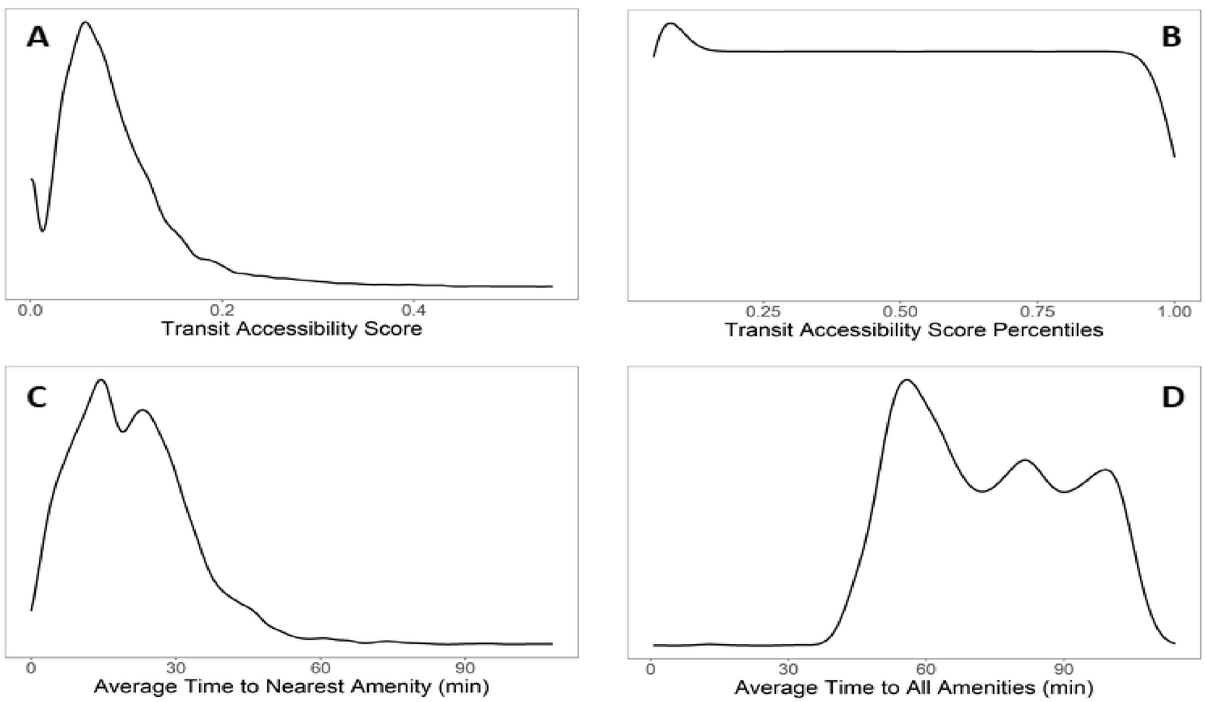


Figure 1: Overall distribution of Greater Vancouver's transit accessibility measures to cultural amenities which include art galleries, libraries, museums, and theatres. All measures were averaged across each type of amenity. **A.** Transit accessibility scores to the nearest amenity. **B.** Percentiles of A's score distribution. **C.** Average time to the nearest amenity (local accessibility). **D.** Average time to all accessible amenities (inter-city accessibility).

The final accessibility measure was time based offering the greatest interpretability for absolute transit times going out from any particular city block. Figure **1C** shows localized access to amenities while figure **1D** shows the relative inter-city access to amenities. Which measure is more useful depends on the amenity and context of urban development projects. For example, access to the nearest hospital is more important than access to all hospitals while access to many art galleries would be more useful than access to a single one. The computed transit times suggest the nearest library, art gallery, museum, and theatre can be reached within 30 minutes for 80% of blocks in Vancouver. On the other hand, 70% of blocks require an average time above 60 minutes to reach any amenity. This is not a particularly good result for inter-city travel, even with the optimistic assumptions of no traffic, no delays, no missed busses, no wrong busses. To understand why the transit network doesn't perform well when considering all amenities, we examined how transit connects city subdivisions in sections 4.2, 4.3 and 4.6.

4.2 Accessibility Measures across Vancouver Subdivisions

Accessibility measure densities of the 12 largest subdivisions of Vancouver were visualized as violin plots ordered by mean at the top. Additional pointers are provided to show how their order changes as different measures are selected. What's interesting is how Maple Ridge and West Vancouver rise from positions 11, 12 (time measure) to positions 5, 6 when considering a score measure.

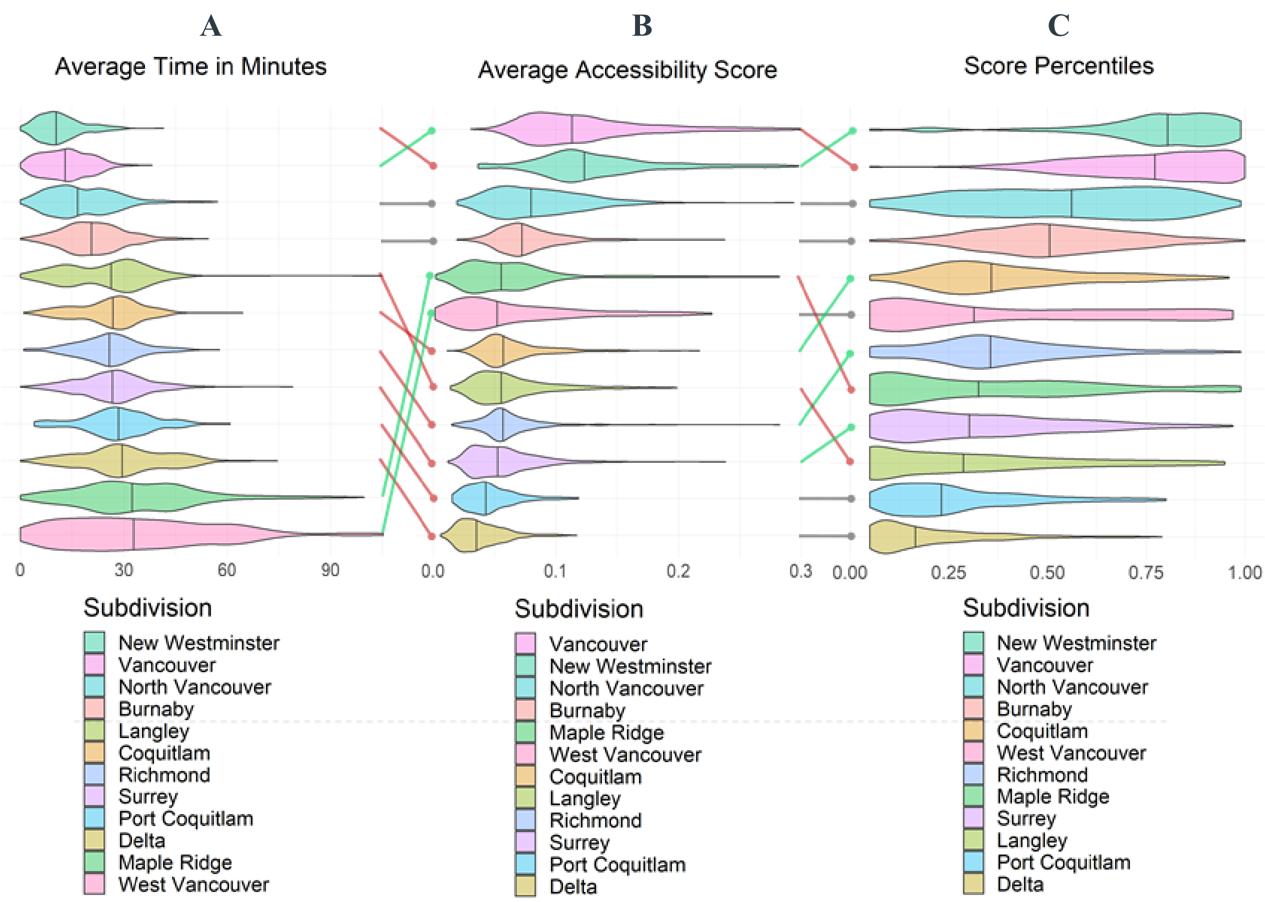


Figure 2. Violin density plot comparing the three transit accessibility measures to the nearest cultural amenity across Greater Vancouver subdivisions. The median value is marked in each density and they're ordered with the highest mean scores on top.

Upon closer inspection, we see that West Vancouver's density presses up against zero in Figure **2A** indicating a higher presence of city blocks that neighbour or occupy cultural amenities (travel times of 0 minutes). When a time of 1-5 minutes converts to a score it becomes 0.2-1, and we can see that in **2B**. This skews scores raising their means higher than other subdivisions ultimately causing them to increase in ranking despite having slightly lower median scores. Subfigure figure **2C** describes accessibility representation between subdivisions giving yet a clearer picture of the relative differences between subdivisions. Ideally, a city should have all violin plots in **2C** be uniform so that each subdivision has similar accessibility representation. However, this is exceedingly difficult in larger cities as it simply takes longer to move from one subdivision to another unless a high speed traffic uninterrupted transit network is adapted to connect these regions, perhaps via more transit train systems. As a result, the subdivisions on the outskirts of metropolitan Vancouver consistently fall near the bottom of these rankings.

4.3 Accessibility Measures over Time

As the travel time matrix has been computed, there are many questions we can answer based on the outcome. One important question is how the accessibility to cultural and art amenities via public transit of each municipality in Metro Vancouver differs from one another and how it changes along the day of the week and the time of the day. In order to view the outcome in a more intuitive way, the average travel time of each subdivision to the nearest amenity, in this case, one of the 4 types (library, gallery, museum, and theatre) can be plotted for every hour from 7am to 7pm. As shown in Figure 3, subdivisions are indicated by contour lines with different colours, and the size of the contour line indicates the total population of the corresponding area. The legend on the right side of the plots presents the areas in a descending order of average travel time. Among the 15 municipalities, Vancouver and New Westminster have the shortest average transit times to the nearest cultural amenity via public transit, which are approximately 20 minutes. Burnaby, North Vancouver, and Port Moody form the group that has the second shortest travel times. The average transit times for the majority are in the range of 35 to 45 minutes. By contrast, people living in Delta need the longest transit time to reach the nearest amenity, especially on the weekends, it takes more than 50 minutes to get to the destination. Port Coquitlam has the second longest average transit time among all cities.

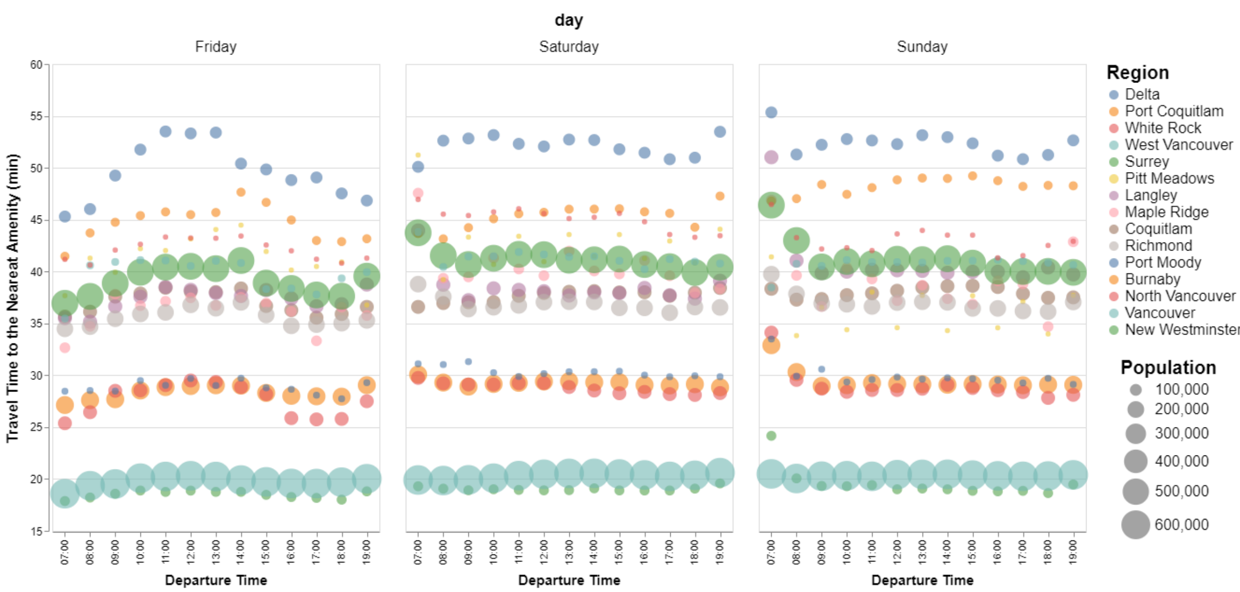


Figure 3. Average Travel Time (minutes) from each subdivision to the nearest cultural Amenity via public transit from 7am to 7pm for weekday (Friday), and weekends (Saturday and Sunday).

On the other hand, Translink, Metro Vancouver's transportation network, schedules more runs during the peak hours. Peak periods on weekdays are 5:30am to 9:30am, and 3pm to 6:30pm (Translink, 2020). Schedules on weekends usually do not have differences during the day. However, weekend transit service starts later in the morning, and finishes earlier in the evening (Translink, 2020). By comparing weekday (Friday), and weekends (Saturday, Sunday) in figure 3, overall weekday has slightly shorter travel time, while Sunday has the longest average travel time for most areas, which matches with the Translink's services that Sunday has the sparsest schedules. Furthermore, according to the first plot in figure 3, all areas have relatively short average transit time during the peak hours on weekdays, since the services are more frequent during these two periods. Similarly, the average transit times to the nearest amenity in the early morning and late afternoon of weekends are longer compared to other time periods for many areas, such as Surrey, Burnaby, North Vancouver, Richmond, and White Rock, which is also consistent with the transit service schedules.

With regards to the population, in the perspective of public transit design, areas with larger populations should have higher public transit accessibility. Overall, the average transit times are consistent with population distribution in Metro Vancouver as areas with larger populations usually have higher accessibilities and vice versa. From the dissemination block data, the populations of City of Vancouver and City of Surrey account for 46.7% of the total population in Metro Vancouver, where Vancouver has 631,486 residents, and Surry has 517,887 residents respectively based on the 2016 census. According to figure x, Vancouver has the shortest average travel time (approx. 20min), while residents in Surry need to spend more than 35 minutes to get to the nearest amenity by public transit. This number is even higher during the weekends. On Sunday morning, on average the trip would take more than 45 minutes. The main reasons could be the physical location of Surry, and the low number of cultural amenities there, which will be discussed in the later sections.

4.4 Visualizations

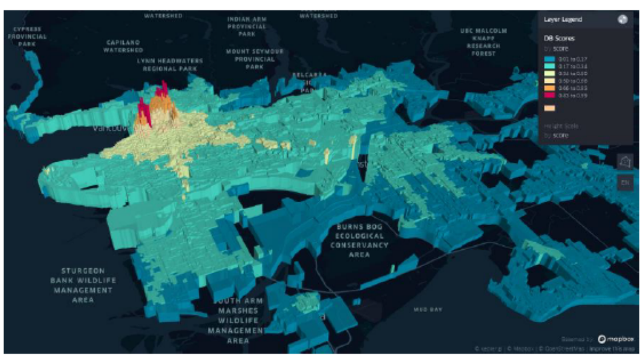
4.4.1 UrbanAccess with Pandana

Figure 4 represents the overall public transit network and the integrated transit and pedestrian network of Metro Vancouver computed using urbanAccess. In the transit network, there are straight lines across southeast Richmond, Delta, and Surrey, which are not roads nor skytrain lines. The transit network containing these lines is due to the lack of transit services in these areas to connect to the other cities, which means these areas are relatively isolated compared to others. UrbanAccess instead computes these edges based on the euclidean distances. Another thing needs to pay attention to is that cities, such as Delta, Langley, and Maple Ridge, have very sparse transit networks compared to Vancouver. This would negatively affect the transit accessibilities in these areas, which will be discussed in later sections. Moverover, the integrated transit and pedestrian network indicates that Vancouver has the most public transit routes, whereas residents in cities in the east of Metro Vancouver only have access to the pedestrian network.

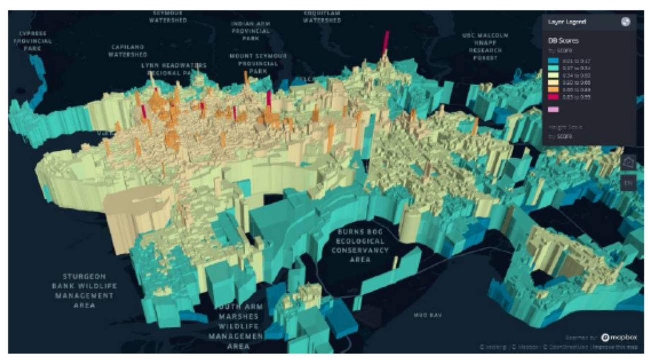
Figure 4. The public transit network (Left) and the integrated transit and pedestrian network (Right) of Metro Vancouver.

4.4.2 Score Visualization in Kepler

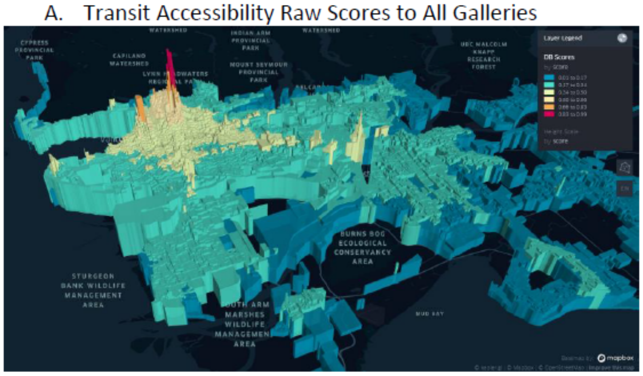
Figure 5 shows the raw scores from each block to cultural amenities. Figure **5A** indicates that most blocks fall into the range of 0.01 to 0.17 (dark blue), and 0.17 to 0.34 (light blue). Only blocks in the City of Vancouver obtain scores more than 0.5. There are extremes (red peaks) in the Downtown Vancouver Area, which have very high scores (>0.83). In Figure **5C** and **5D**, the scores to all museums, and all theatres from the blocks have very similar geographic distribution as to all galleries. By contrast, the score distribution in Figure **5B** to all libraries tells very different stories. Most outskirts have scores in the range of 0.17 to 0.34, while the majority fall into the range of 0.34 to 0.65. The differences in scores are much smaller than other types, which implies that the accessibility scores to libraries are relatively even and within the middle range in the whole area.



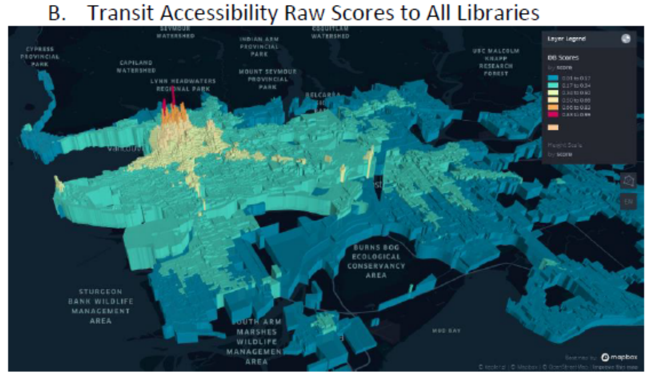
A. Transit Accessibility Raw Scores to All Galleries



B. Transit Accessibility Raw Scores to All Libraries



C. Transit Accessibility Raw Scores to All Museums



D. Transit Accessibility Raw Scores to All Theatres

Figure 5. Kepler maps - raw accessibility scores from each dissemination block to all amenities (**A**. galleries, **B**. libraries, **C**. museums, and **D**. theatres).

4.4.3 Transit Time Visualization in Kepler

Kepler.gl is a convenient tool for comparing multiple sets of large-scale spatial data. It can present two scrolling-linked maps in a dual view making it easy to inspect for granular differences. Users can select time windows and also plot additional layers of data on top of the existing map such as the amenity points. Figure 6 shows a screenshot of an isochrone comparison to the nearest gallery between Friday and Saturday. The travel time to the nearest cultural amenity via public transit is more or less similar across different days of the week with the most differences being in the least accessible blocks. Those cases are highlighted with green circles.

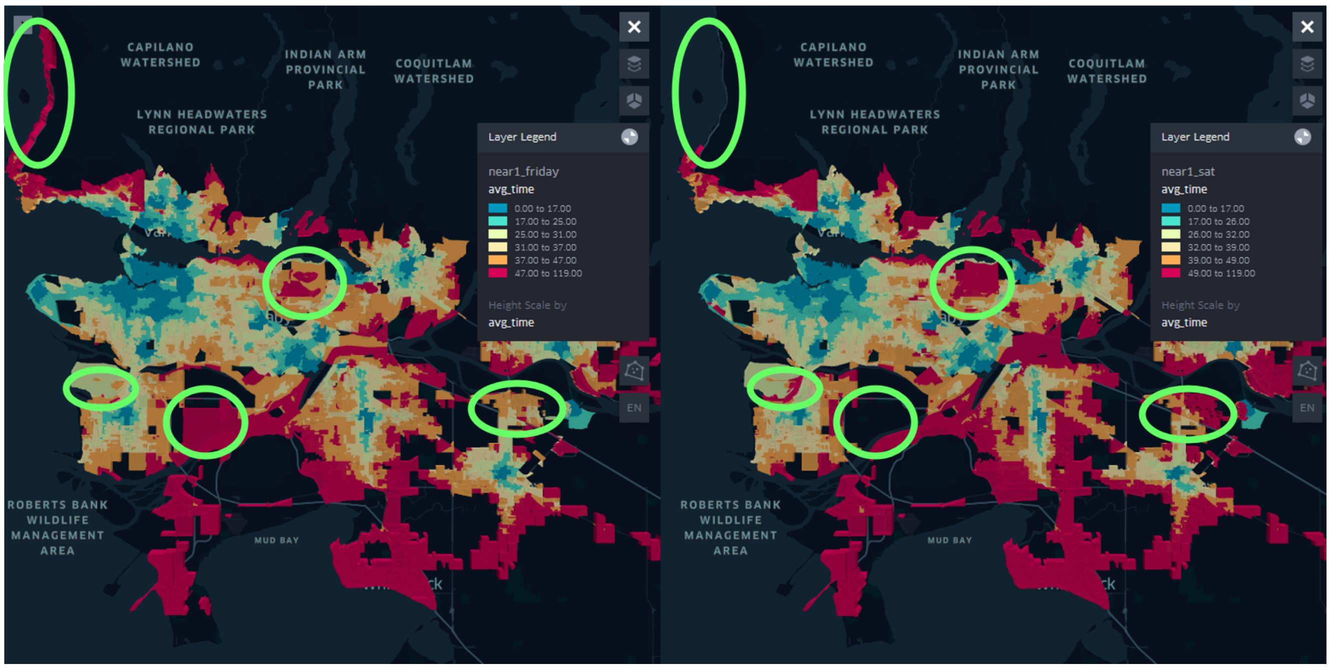


Figure 6. Kepler isochrone maps for transit times from each dissemination block to the nearest gallery from 7am to 8:30am on Friday (Left), and Saturday (Right). Differences shown with green circles.

The main differences are shown in areas which are distant from the City of Vancouver. For example, on Friday it takes more than 49 minutes to take transit from blocks in Horseshoe Bay (left upper corner) to the nearest gallery, while on Saturday, there is no public transit access directly from this area to any gallery. Same thing happens for the blocks in southeast Richmond. For Burnaby Mountain, as well as many blocks in Surrey and Langley, the transit time increases from the range of 39 to 49 minutes to more than 49 minutes. Translink's weekend services have a strong negative impact on the accessibilities of blocks that are distant from Vancouver, whereas blocks in Vancouver usually have higher accessibilities to any type of cultural amenities on both weekdays and weekends.

Furthermore, locations of the cultural and art amenities also affect the transit accessibilities. Small yellow dots in figure 7 represent the locations of all galleries in Metro Vancouver. The type of amenities can be selected based on users' interests for both layers. A large number of galleries concentrate in Downtown Vancouver, the areas close to downtown, UBC, as well as New Westminster. Therefore, transit times for blocks in these areas are shorter, which are in the range of 0 to 17 minutes. By contrast, areas that are shown in red do not have any galleries that are close to. The trips from these blocks usually take more than 47 minutes to get to the nearest gallery. Overall, areas with more cultural amenities have higher transit accessibilities, while people from areas with limited amenities need longer time to cross the areas by public transit.

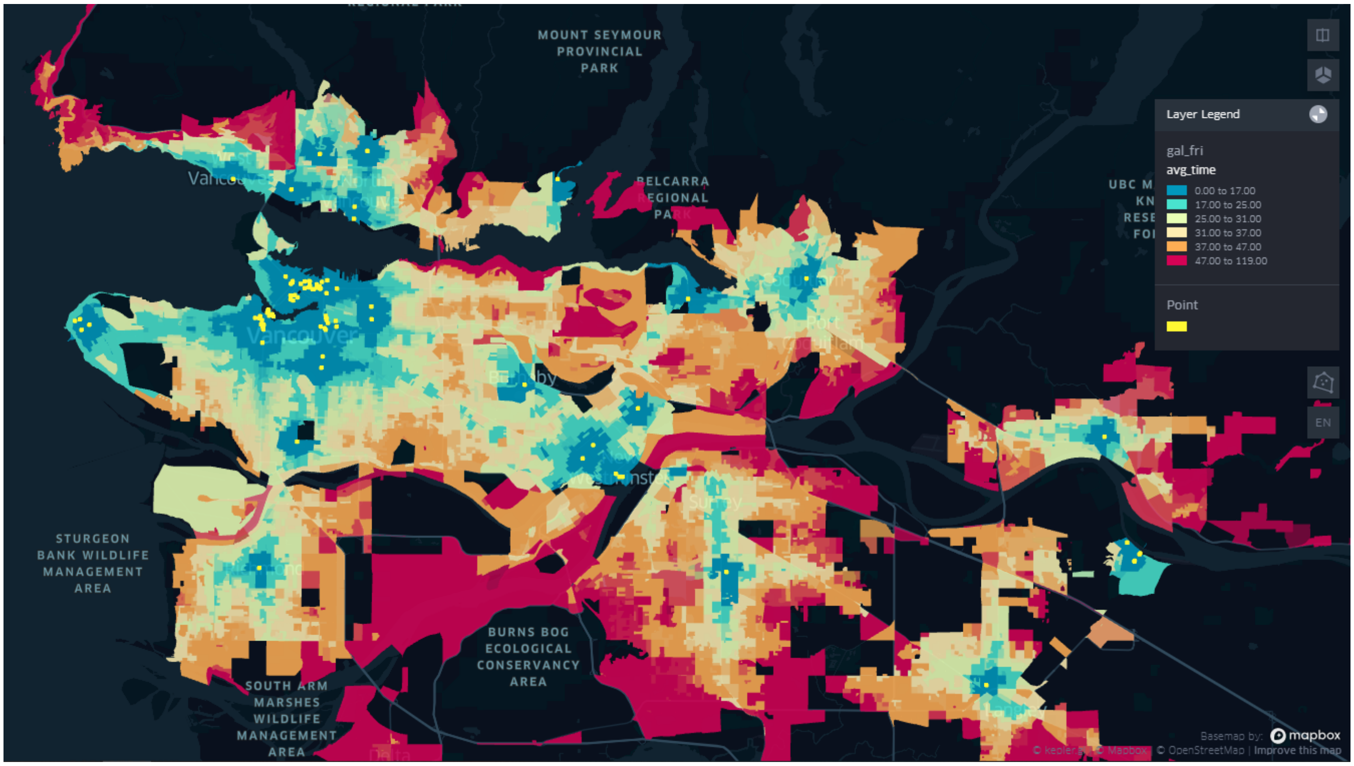


Figure 7. Kepler isochrone to the nearest gallery on Friday with the locations of all galleries (yellow dots).

4.4.4 Score Visualization in Leaflet

In score visualization maps (figure 8), the colour of blocks is based on the percentile of raw accessibility scores. It uses a discontinuous colour scheme from red to green where each colour represents a 20th percentile chunk of scores. Green corresponds to the top 20th percentile of score (highest transit access). These green zones tend to be found where transit is most developed, where population is most dense, and where amenity density is greatest. This would include downtown Vancouver, New Westminster, and Richmond. On the other hand, city extremities fall in the bottom 20th percent of scores. Curiously, despite representing only 20% of the data, it seems almost half of the map is red. This is because blocks of extremities tend to occupy more land giving the impression that there's more “bad scores” than in reality. Furthermore, large blocks can bias the routing engine because it may consider the time required to walk from the center of these large plots of land to the nearest road which inevitably worsens the travel time. When comparing the four amenities, it is clear that libraries are found in most local communities while theatres, galleries, and museums tend to be more centralized to the largest urban area such as downtown Vancouver, New Westminster, and Richmod.

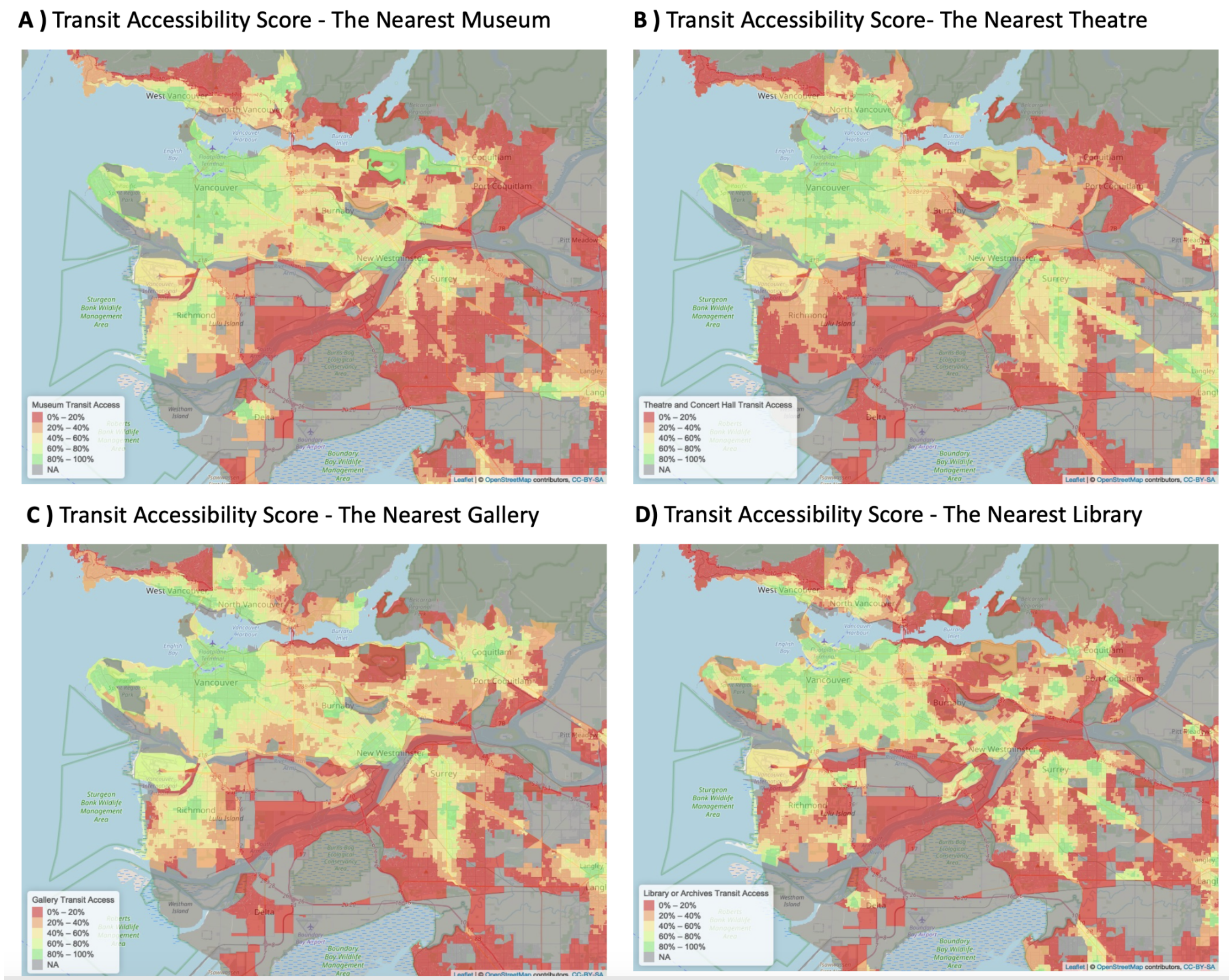


Figure 8. Score Visualizations for four types of amenity. **A)** percentile of accessibility score of the nearest museum with weights, **B)** percentile of accessibility score of the nearest theatre with weights, **C)** percentile of accessibility score of the nearest gallery with weights **D)** percentile of accessibility score of the nearest library with weights.

4.4.5 Isochrone Visualization in Leaflet

The isochrone maps use a similar divergent green - red colour scheme but are far more interpretable since the absolute time in minutes to the nearest amenity is displayed. When considering libraries versus the other amenities, similar patterns to the score measures are seen where libraries are a much more common amenity than the others. What is interesting is how between **9B** (Theatre) and **9C** (Gallery), Coquitlam (top right), changes from orange/red to yellow green simply due to the presence of amenities. This is important to note since the transit system is identical in both maps, yet the measure is penalizing the transit network simply due to the absence of amenities, and not actually due to the inefficiency of transit. As such, if the primary goal is to improve the transit network, an aggregate across all amenities should be considered when computing transit accessibility measures.

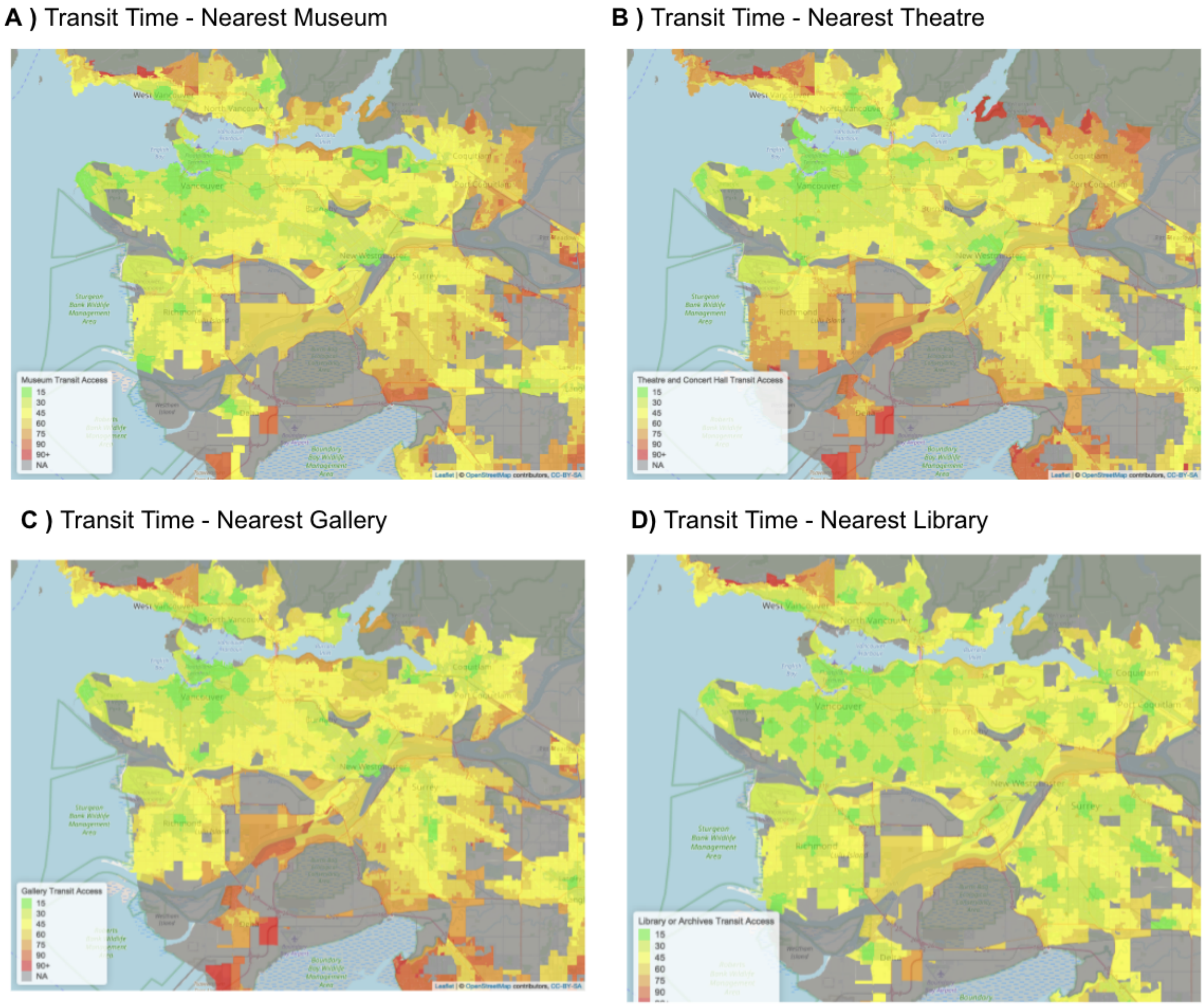


Figure 9. Isochrone Visualizations. **A)** Isochrone map using actual transit time to the nearest museum, **B)** Isochrone map using actual transit time to the nearest theatre, **C)** Isochrone map using actual transit time to the nearest gallery, **D)** Isochrone map using actual transit time to the nearest library.

4.5 The Significance of Weighing Amenities

In the context of network analysis, it can be useful to consider the popularity of specific amenities if the focus is more on amenity accessibility than transit accessibility. For instance, if analyzing urban food accessibility, one would weigh convenience stores far less than grocery stores since they provide far less nutrition and variety than a grocery store might. In the same way, individuals may prefer to visit one museum or art exhibit over another due to its popularity. Thus, weights can offer a very realistic correction for amenity significance.

As explained in 3.4 of methodology, weights for libraries and other cultural institutions were obtained and incorporated into the score computation. A two-sided Welch sample t-test, a Wilcoxon rank sum test, and a Kolmogorov-Smirnov test were performed to assess for significant differences between the weighted and unweighted score distributions. All tests returned a definite significance with p-values less than 1e-16 and a mean difference of 0.047 - 0.071 regarding the score measure. However, since the weights were simply added to the numerator of the score (3.3 equation 1), it is not surprising that the difference is significant as the entire distribution would have been shifted. The statistical tests and distribution are visualized figures 17 and 18 of the appendix.

Despite the significance, an average shift of 4.7% - 7.1% does not confirm the weights had a true effect on modifying the score distribution since weighted and unweighted distributions were almost identical in shape. When visualized on a map, the differences remained sparse since suggesting no perceptible change in the data. The weights may not have been large enough to modify the distribution since they had a scale of 0-1 while travel time in the denominator had a scale of roughly 10-60. Secondly, the values of the weights were too similar (most being around 0.4 to 0.6) resulting in very minimal impacts from block to block. Thirdly, the hours and day open feature may be redundant reducing the weight of other features. Lastly, the overall weights of the features need to be considered relative to other amenities (ie. Science World occupies 6% of total reviews while Delta museum occupies only 0.01%) to achieve more variation in values. In either case, future work may want to consider exploring alternative methods to incorporating such amenity weights to see more impact.

4.6 Unsupervised Analysis

Public transit is a key factor to consider when residents choose where to live, but also, having cultural and art amenities may make communities more appealing as well. In the case of Metro Vancouver, features related to these two factors may have very different influences on different subdivisions. To investigate this question, k-means clustering was conducted for all cities in Metro Vancouver across seven variables: the average accessibility score, the total number of all four types of cultural amenities, the total population, the average transit time to the nearest amenity, the total number of bus stops, the average number of buses passing each stop, and the average proximity scores to public transit for each city. The last proximity scores were obtained from a study earlier this year by Statistics Canada (Alasia *et al*, 2021), where the measures are based on walking and driving distances. The k-means analysis suggested five clusters among all subdivisions, which are indicated in figure 10.

Among the five clusters, the City of Vancouver forms its own cluster, and it is distant from the other four clusters in the plot. Regarding public transit accessibility to cultural amenities, it has outstanding performances on most variables, including accessibility score, the transit time to any types of amenities, total number of all types of amenities, as well as total number of bus stops. On the other hand, New Westminster also forms its own cluster far from the others. One strong variable which differentiates it is the total number of cultural amenities. Vancouver has the most at 195, while New Westminster has the second most at 80. By contrast, all other cities have less than 22. Furthermore, Burnaby, North Vancouver, Surrey, and Richmond form the third cluster, which is the closest one to the City of Vancouver. Geographically, three of these cities share a border with the City of

Vancouver, while Surrey shares a border with New Westminster. The yellowish-green cluster (cluster 2) contains Delta, Coquitlam, and Port Coquitlam. None of these areas are directly connected to Vancouver as there is always another city in between. They are essentially the second outer circle from Vancouver. The blue cluster contains the last five cities: West Vancouver, Langley, Pitt Meadows, Maple Ridge, and White Rock. These five cities are the furthest areas from the City of Vancouver. They are on the outermost circle of the entire metropolitan area. Hence, the public transit accessibility to cultural amenities is heavily influenced by the geographic location of the city. As the City of Vancouver is the centralized region, the accessibility decreases progressively as the location gets far and far from Vancouver. The main reason is that Vancouver has the most convenient transit services and most cultural amenities. In addition, the closeness to amenities also decreases the travel times by public transit.

An interesting point when observing a map of metro Vancouver is that boundaries of these clusters are typically physical barriers. For example, Fraser River separates Vancouver (pink cluster) and Richmond (red cluster), and Vancouver Harbour separates Vancouver and West Vancouver (blue cluster). This may suggest that transit services are not sufficiently connecting these areas across physical barriers.

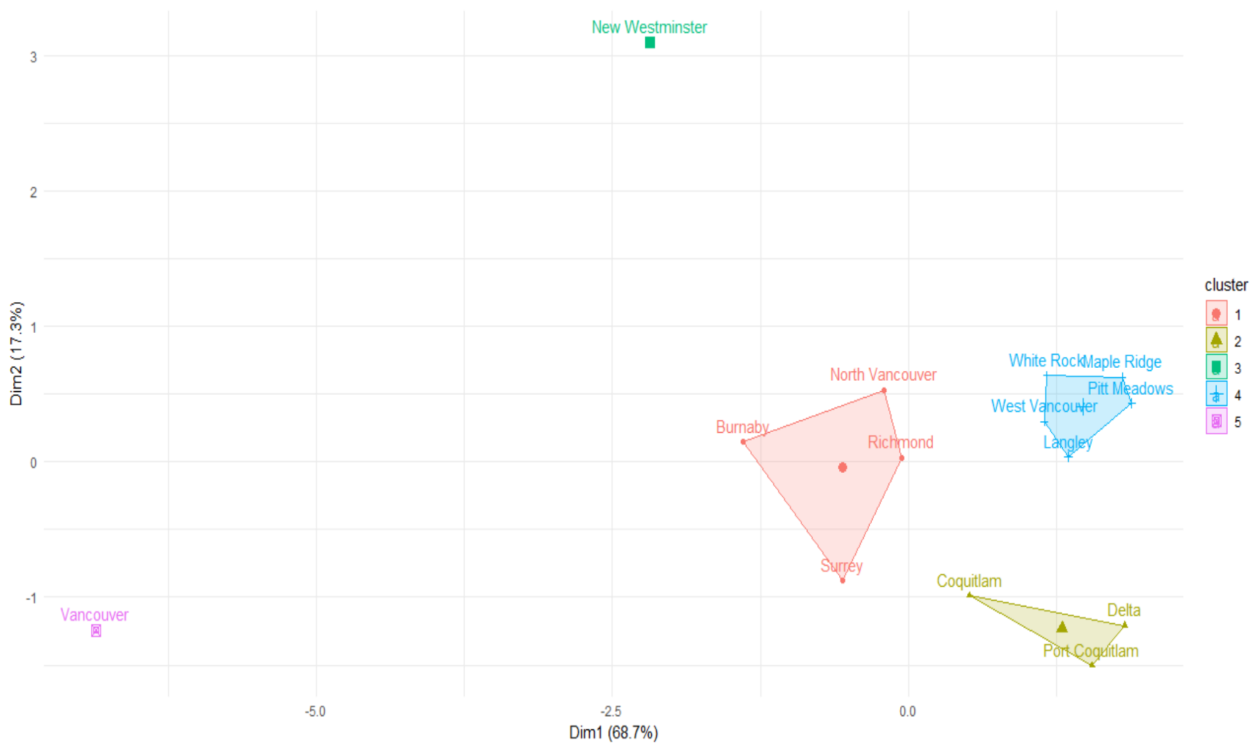
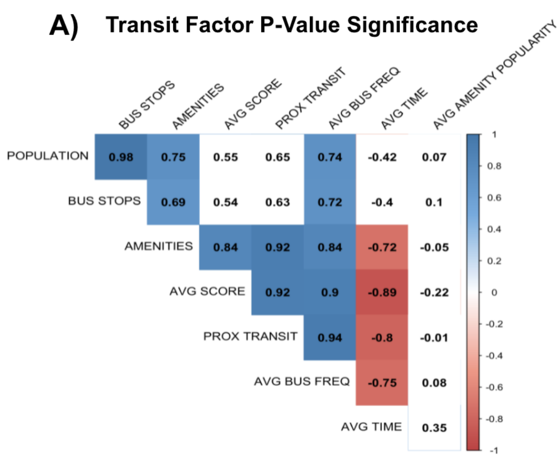


Figure 10. K-means clustering results for all cities in Metro Vancouver. Variables: average accessibility scores, total number of cultural amenities, total population, average transit time to the nearest amenity, total number of bus stops, bus frequency, and proximity scores to public transit.

4.7 Correlation Analysis

In addition, we want to investigate the correlation between the eight variables that were used in clustering analysis. Figure **11A** shows the correlations among variables where only significant correlations ($p < 0.01$ using a Perarson t-test) are coloured in.



B)

```
	Pearson's product-moment correlation

data:  df_pca$`BUS STOPS` and df_pca$`AVG SCORE`
t = 2.2438, df = 12, p-value = 0.04449
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.01835985 0.83373532
sample estimates:
      cor
0.5436435
```

C)

```
	Pearson's product-moment correlation

data:  df_pca$`BUS STOPS` and df_pca$`AVG TIME`
t = -1.522, df = 12, p-value = 0.1539
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.7687602  0.1631423
sample estimates:
       cor
-0.4022565
```

Figure 11. A) Transit Factor correlation table using matrix of the p-value of the correlation. **B)** Pearson's correlation test for the number of bus stops v.s. accessibility score **C)** Pearson's correlation test for the number of bus stops v.s. transit time. **Note:** in A), a colored score indicates whether the correlation coefficient is significant at level of 0.01. Details on variables are in table 4 of the appendix.

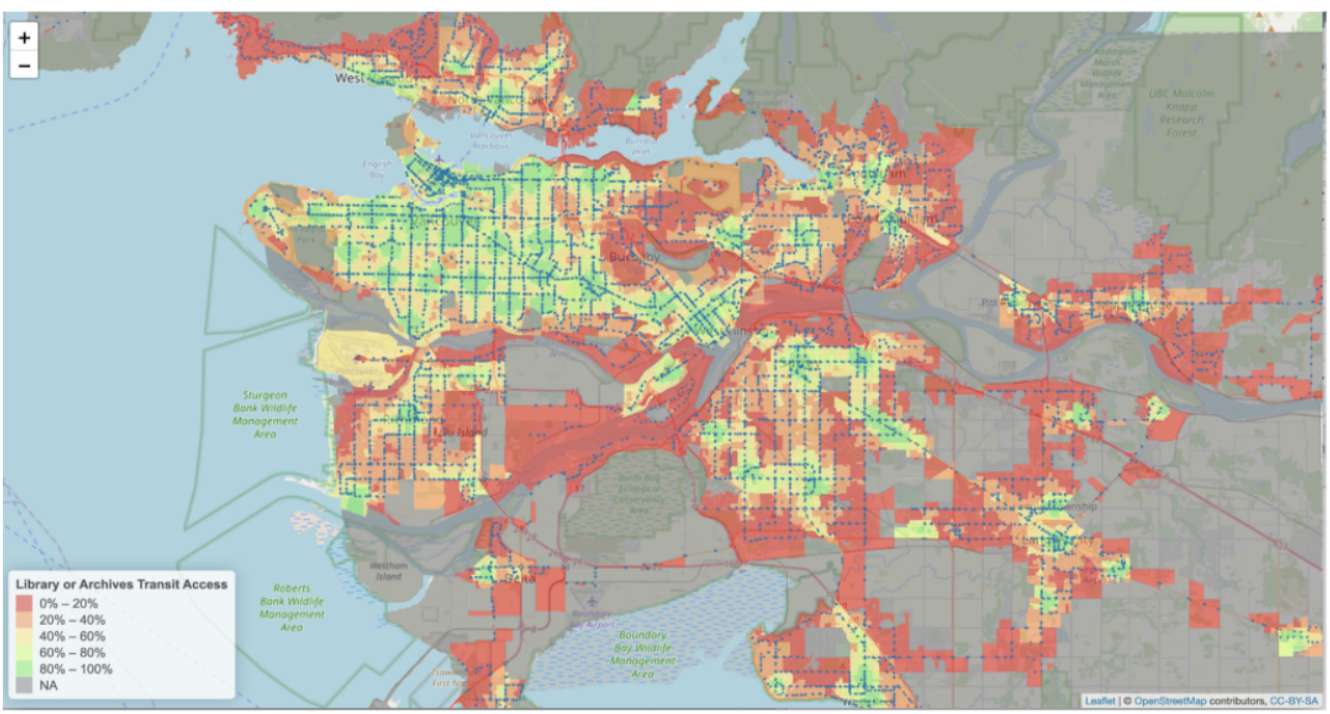


Figure 12. Bus stop network overlaid on the accessibility scores to the nearest 1 library or archives.

Intuitively the number of bus stops should have a strong negative effect on transit access; however, **11B** shows that the number of bus stops are insignificantly correlated with both accessibility scores (p = 0.01) and transit time (p = 0.15). Having more bus stops to stop at should only slow down transit operations but this is not the case so a significance of 0.01. The greatest correlating factor for reducing trip time and improving accessibility scores is bus frequency, so a possible way to address underrepresented regions in Vancouver would be through increased bus frequency, not bus stops. We can visualize the locations of bus stops as seen in figure 12. Despite having substantial bus stops in orange and red regions, it's likely the frequency and lack of amenities that causes the score to be so poor.

4.8 Efficiency Measures

Even if one has the transit accessibility scores, it does not indicate whether the level of transit meets the needs of each block. If a particular block requires a large amount of transit it is important that it has access to that transit. However if the block does not require a large amount of transit then it would be a waste of resources to operate transit systems where they would not be utilized. This is how we define efficiency, the difference between accessibility scores and needs score as summarized by equation 7. Percentiles are taken for two reasons. Firstly, so that the final efficiency score is between -1 and +1, with 0 indicating highest efficiency. Secondly, to ensure both variables have the same distribution and can be comparable. That way, we can match the *y*th percentile accessibility scores with the *y*th percentile needs scores.

$$Eff = \%ile(Transit\ Accessibility\ Score) - \%ile(Transit\ Needs\ Score)$$

Equation 7

The needs of each block are defined as the mean of three parameters: the population, traffic intensity and the amenity density. The population reflects the base accessibility requirements of the block, but a better measure for future work would be the percentage of population depending on transit. Traffic intensity restricts a person's access to the city by increasing travel time. This can be remediated by reducing cars on the road with more transit, thus serving as an important variable for computing the needs. The amenity density is the number of destinations in proximity to each block that people can access. With a larger density, more people will want to access the block, increasing the required need for accessibility.

A diverging color scheme of red, white, and blue -- white representing the inner 40th percentile of values -- was used to emphasize blocks on the extremes of the efficiency scoreset. Blocks shown in red represent a lack of accessibility while blocks shown in blue indicate an excess of accessibility. White was used to indicate highly efficient blocks whose needs are similar to their level of accessibility. The color percentiles are detailed in Table 5 of the appendix.

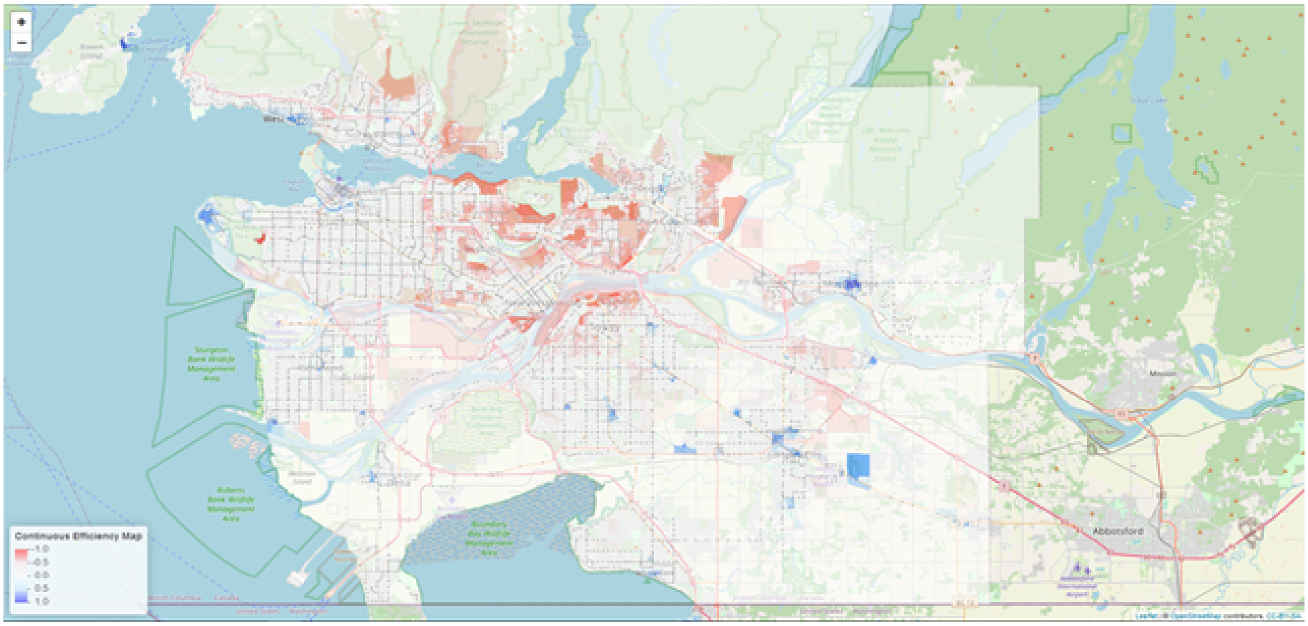


Figure 13. Efficiency scores displayed using a continuous color scheme. A quantitative color scheme also exists but is in the appendix as inadvertently altered the overall efficiency of the map due to varying block size. Blocks shown in red have a low efficiency and blue for blocks with a higher efficiency.

Figure 14. Distribution of efficiency scores. The score forms a distribution that is slightly shifted towards the lower efficiency.

Overall the Greater Vancouver Area has a mean efficiency score that is slightly biased towards the lack of accessibility with 57.9% of the score being less than 0. For each city within the Greater Vancouver Area only three municipalities, West Vancouver, Maple Ridge, Langley, had a median efficiency greater than 0, while three other municipalities, North Vancouver, Vancouver and Richmond, have an efficiency near 0, figure 14 & 15. Six other municipalities including the Greater Vancouver Area overall have a median score below 0. This indicates that overall the Vancouver area is in need of increased accessibility particularly in the Burnaby and Coquitlam area with the efficiency score distributions being focussed slightly towards -1. Municipalities on the fringes of the city such as Delta and Langley have more pointed/concentrated distributions. These distributions are associated with cities with smaller transit networks with few amenities relying on a few bus routes causing the overall density to be concentrated around specific efficiency scores. In contrast cities with large transit networks like Surrey and North Vancouver have a broad range of efficiency scores.

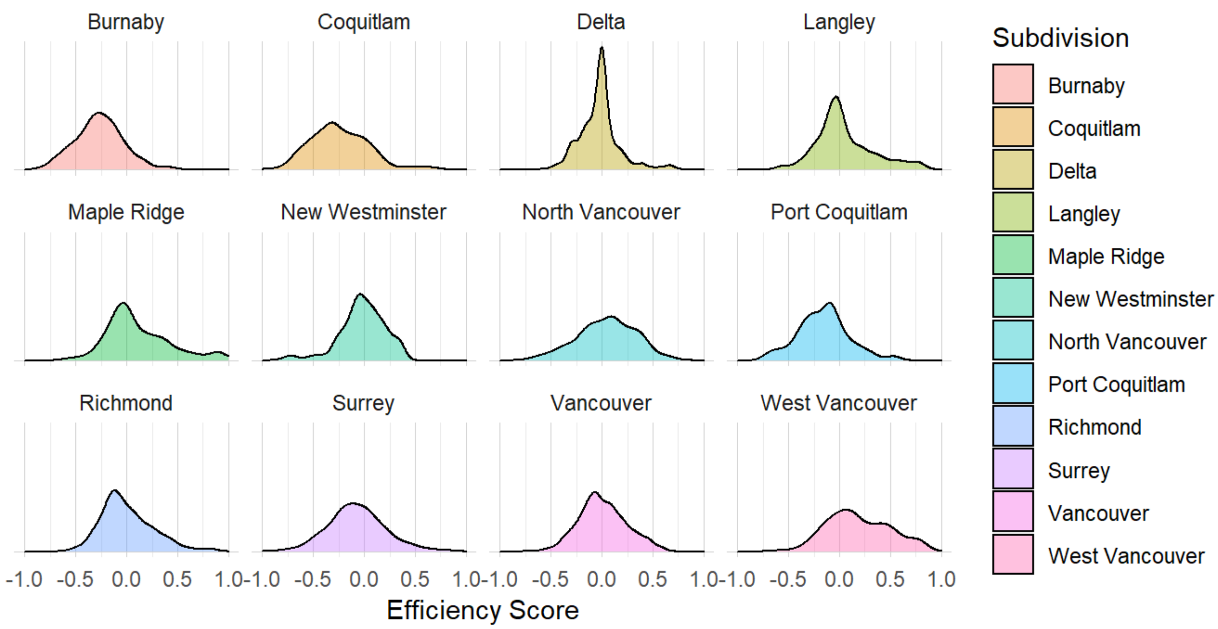


Figure 15. Density plots of the efficiency scores grouped by Greater Vancouver Area subdivisions.

Overall blocks with lower efficiencies occur in the NE quadrant of the Greater Vancouver Area in the cities of Burnaby, Coquitlam and Surrey. These cities have large populations and traffic congestion with very low accessibility. What is of particular interest is the boundary between the cities of Vancouver and Burnaby called “Boundary Rd”. Blocks to the west in Vancouver overall vary between an efficiency score of 0.1 to -0.3. However to the east of the road, you see a similar trend but with the appearance of blocks with a large population but with lower accessibility scores. This suggests the presence of gaps in the transit network that are not servicing areas within the city to accommodate different communities and would suggest a difference in the strategies between the different cities on the layout of the transit networks.

The efficiency score however is only a show of concept and more data is required to increase its accuracy. In particular, the efficiency uses traffic survey data to estimate the traffic in specific areas which lack accuracy, are infrequent and do not account for the size or type of roadway. As such it is difficult to precisely measure congestions throughout the network. The efficiency data also accounts for every member of the population rather than only members of the population who actually use transit as their main source of transportation.

5 Limitations and Assumptions

5.1 Limitations of Travel Times

The travels calculated from R5 are optimal travel times that do not include traffic volumes, delays, missed busses, and roadwork or construction to name a few. It also fails to consider seasonal variations in overall levels of congestion in a city. A more comprehensive but also more difficult and time intensive approach for realistic travel times would be to directly work with transit GPS information since most are already generating such data. Access to these databases could allow for more realistic calculations of actual travel times and uncertainties than a simulated transit network.

5.2 Sparse Traffic Data

The traffic correction used for calculating the need of the efficiency score was calculated based on the normalized mean of all vehicle counts done at specific monitoring sites across the city of Vancouver. To account for the variations in the density of monitoring sites across the city, the mean score of the blocks was calculated by taking the mean of all the vehicle counts within a 5 km radius of the center of each block. By calculating the mean count based on the Euclidean distance the connectivity of the network is not accounted for. If a monitoring site is located along a highway, neighboring blocks would have a higher vehicle count even if there is not any direct access to the highway. To account for this, the traffic counts should be based on the nearest monitoring sites based on the road network using R5.

6. Conclusions and Future Work

To create an open source scalable framework for computing transit accessibility measures, Conveyal's R5 algorithm was leveraged to efficiently compute shortest path travel times for over 5.3 million transit routes (15,197 origins * 353 destinations), 360 times a day over three days. This equates to roughly 5.7 billion shortest path searches and was accomplished in under an hour for one of Canada's largest metropolitan areas, showing that other metropolitan areas in Canada can be similarly addressed without large computational hurdles. Furthermore, we detail the advantages, disadvantages, and possible use cases of three unique accessibility measures which are calculated from the travel time matrix. These include time based, score based, and percentile based accessibility measures. Lastly, we illustrate how these measures can be used for analysis both visually on choropleth maps, and statistically using various techniques such as k-means clustering and Pearson's correlation coefficient matrices. We were also able to leverage these accessibility measures to probe for new ways of understanding network efficiency, and developed an experimental *needs* metric to better visualize underrepresented regions of Vancouver.

Regarding our case study on Metropolitan Vancouver, we were able to identify areas of both high and low cultural transit accessibility not only across regions but also across hours of the day, days of the week, and amenity types all wrapped in a deployed dashboard. Specifically, Vancouver fares relatively well offering 80% of dissemination blocks a maximum of 30 minutes to the nearest cultural amenities. However, intercity connectedness was lacking, requiring 70% of dissemination blocks anywhere between 1 and 2 hours to reach any cultural amenity of interest. We also confirm that bus frequency is a far better indicator of transit access than simply the number of bus stops, which may be a consideration when addressing underrepresented regions in the city, particularly those which score low on efficiency measures.

In any case, future work should consider a few key points. If the aim is performance, other routing engines, like GraphHopper, should be considered for computing the travel time matrices as slightly different backend algorithms may yield further benefits to such frameworks. If the aim is to further build on existing metrics for network analysis, whether it be for equity analysis (how are demographics represented by transit accessibility measures) or sustainability analysis (how efficient is the transit network), the retrieval of more data will be required. Specifically data regarding traffic and transit users by dissemination block. Lastly, if realism is the objective for future work, then sampling existing travel time data from transit mounted GPS may help to identify bias in these “ideal trip” routing engines. Such bias can include traffic choke points during particular times of day and consistently late busses. The routing engine results can then be corrected for added realism and validity.

We conclude by providing a first iteration for addressing the lack of transit accessibility measures, by providing a high performing framework for converting basic -- and readily available -- GTFS network data into such measures. Most importantly, this framework can be scaled across new amenity types (schools, hospitals, parks, grocery stores, etc.) and across other metropolitan areas in Canada to provide a basic methodology for future Census studies regarding transit accessibility measures. This framework provides a basis for future analysis and development that can be adopted by city planners and policy makers to ensure equitable, sustainable, and optimal urban growth for population segments which depend on public transit.

7 References

1. Alasia, A., Newstead, N., Kuchar, J., & Radulescu, M. (2021). Measuring proximity to services and amenities: An experimental set of indicators for neighbourhoods and localities. https://www150.statcan.gc.ca/n1/pub/18-001-x/18-001-x2020001-eng.htm.

2. Albacete, X., Olaru, D., Paül, V., & Biermann, S. (2017). Measuring the accessibility of public transport: A critical comparison between methods in Helsinki. *Applied Spatial Analysis and Policy*, *10*(2), 161-188.

3. Cheng, Y. H., & Chen, S. Y. (2015). Perceived accessibility, mobility, and connectivity of public transportation systems. *Transportation Research Part A: Policy and Practice*, *77*, 386-403.

4. Deboosere, R., & El-Geneidy, A. (2018). Evaluating equity and accessibility to jobs by public transport across Canada. *Journal of Transport Geography*, *73*, 54-63.

5. EMBARQ. (2013). Saving Lives with Sustainable Transport. World Resources Institute.

6. Government of British Columbia. (2015). Traffic Data Program - Legacy Information. https://www.th.gov.bc.ca/trafficdata/legacy.html

7. Handy, S. L., & Niemeier, D. A. (1997). Measuring accessibility: an exploration of issues and alternatives. *Environment and planning A*, *29*(7), 1175-1194.

8. Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of planners*, *25*(2), 73-76.

9. Koenig, J. G. (1980). Indicators of urban accessibility: theory and application. *Transportation*, *9*(2), 145-172.

10. Liu, S., & Zhu, X., (2004). An integrated GIS approach to accessibility analysis. *Trans. GIS 8,* 45–62.

11. Lubitow, A., Rainer, J., & Bassett, S. (2017). Exclusion and vulnerability on public transit: experiences of transit dependent riders in Portland, Oregon. *Mobilities, 12(6),* 924-937, DOI: 10.1080/17450101.2016.1253816

12. Martin, A., Y. Goryakin & M. Suhrcke. (2014). Does active commuting improve psychological wellbeing? Longitudinal evidence from eighteen waves of the British Household Panel Survey. *Preventative Medicine. 69,* 296–303.

13. Mo, B., Cao, Z., Zhang, H., Shen, Y., & Zhao, J. (2021). Competition between shared autonomous vehicles and public transit: A case study in Singapore. Transportation Research Part C: Emerging Technologies, 127, 103058.

14. Morris, J. M., Dumble, P. L., & Wigan, M. R. (1979). Accessibility indicators for transport planning. *Transportation Research Part A: General*, *13*(2), 91-109.

15. Pereira, R. H., Saraiva, M., Herszenhut, D., Braga, C. K. V., & Conway, M. W. (2021). r5r: Rapid realistic routing on multimodal transport networks with r5 in r. Findings, 21262.

16. Roberts, B. (2015). The third industrial revolution: implications for planning cities and regions. Work. Pap. Urban Front, 1.

17. Royal Society for Public Health, Vision, Voice and Practice. (2016). Health in a Hurry: The Impact of Rush Hour Commuting on Our Health and Wellbeing.

18. Saghapour, T., Moridpour, S., & Thompson, R. G. (2016). Public transport accessibility in metropolitan areas: A new approach incorporating population density. *Journal of Transport Geography*, *54*, 273-285.

19. Saif, M. A., Zefreh, M. M., & Torok, A. (2019). Public transport accessibility: a literature review. *Periodica Polytechnica Transportation Engineering*, *47*(1), 36-43.

20. Sallis, J. F., Frank, L. D., Saelens, B. E., & Kraft, M. K. (2004). Active transportation and physical activity: opportunities for collaboration on transportation and public health research. *Transportation Research Part A: Policy and Practice, 38(4),* 249-268.

21. Sheppard, S. (2014). Measuring the economic and social impacts of cultural organizations. *Community Development Investment Review*, *10*(2), 43-48.

22. Statistics Canada. (2017). Census of Population, 2016: Commuters using sustainable transportation in census metropolitan areas. Minister of Industry. Catalogue no. 98-200-X2016029. ISBN 978-0-660-23763-3.

23. Translink. (2020). Managing the Transit Network. https://www.translink.ca/plans-and-projects/strategies-plans-and-guidelines/managing-the-transit-network

24. Wang, C. H., & Chen, N. A. (2015). A GIS-based spatial statistical approach to modeling job accessibility by transportation mode: case study of Columbus, Ohio. *Journal of transport geography*, *45*, 1-11.

25. Washington Post. (2013). Washington: A world apart. https://www.washingtonpost.com/sf/local/2013/11/09/washington-a-world-apart/

8 Appendix

Amenity Type	Amenity Name	Open Days	Open Hours	Ratings	Total Reviews
Gallery	1.01	28.28	30.30	11.11	11.11
Library	1.16	25.58	25.58	19.76	19.76
Museum	4.35	29.35	29.35	6.52	6.52
Theater	4.00	61.33	62.67	8.00	8.00

Table 3. Missing Values from Google Search and Google Place API requests (in %)

```
        Welch Two Sample t-test

data:  no_wt$x and wt$y
t = -9.6271, df = 28583, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.07178031 -0.04749604
sample estimates:
mean of x mean of y
-2.464655 -2.405017


        Wilcoxon rank sum test with continuity correction

data:  no_wt$x and wt$y
W = 95769443, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0

p-value will be approximate in the presence of ties
        Two-sample Kolmogorov-Smirnov test

data:  no_wt$x and wt$y
D = 0.049853, p-value = 7.772e-16
alternative hypothesis: two-sided
```

Figure 16. Welch two sample t-test, Wilcoxon rank sum test, and two-sample Kolmogorov-Smirnov tests on unweighted and weighted score distributions returning significance p < 1e-16.

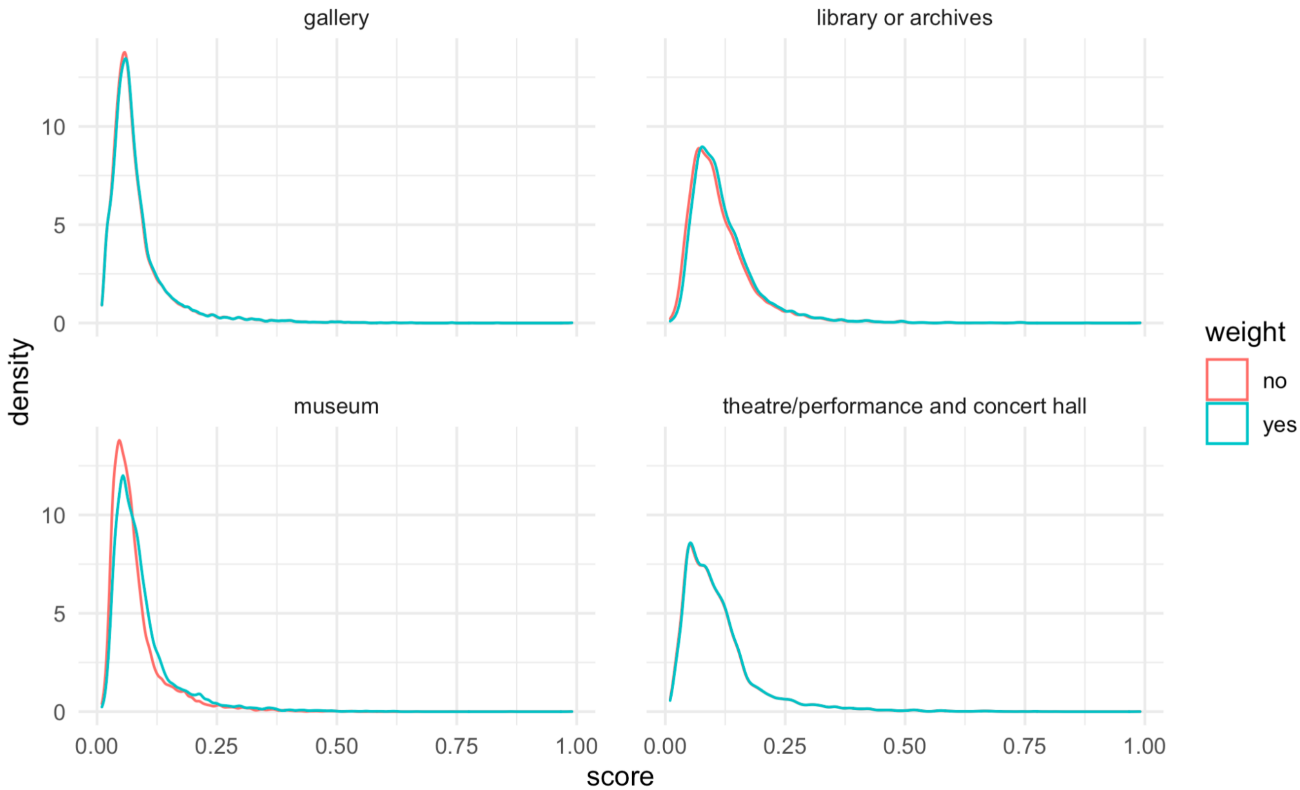


Figure 17. The effect of amenity weights on the score distributions for each amenity type.

Table 4. Transit Variable Name and Explanation

Variable	*Description*
AVG TRANSIT TIME:	The average travel time using public transit from blocks to the nearest amenities by subdivision
AVG SCORE:	The average of accessibility scores by subdivision.
AVG AMENITY POPULARITY:	The average amenity weight by subdivision.
AMENITIES:	The total number of amenities in each subdivision.
BUS STOPS:	Total number of bus stops in each subdivision.
BUS FREQ:	The average of total buses that pass each bus stops in a day by subdivision.
POPULATION:	Total population in each subdivision
PROX TRANSIT:	The average proximity scores to public transit for each city.

Color	*Efficiency Percentile*
Dark Blue	**90 - 100**
Blue	**80 - 90**
Light Blue	**70 - 80**
White	**30 - 70**
Light Red	**20 - 30**
Red	**10 - 20**
Dark Red	**0 - 10**

Table 5. Efficiency Map Percentile Color Scheme

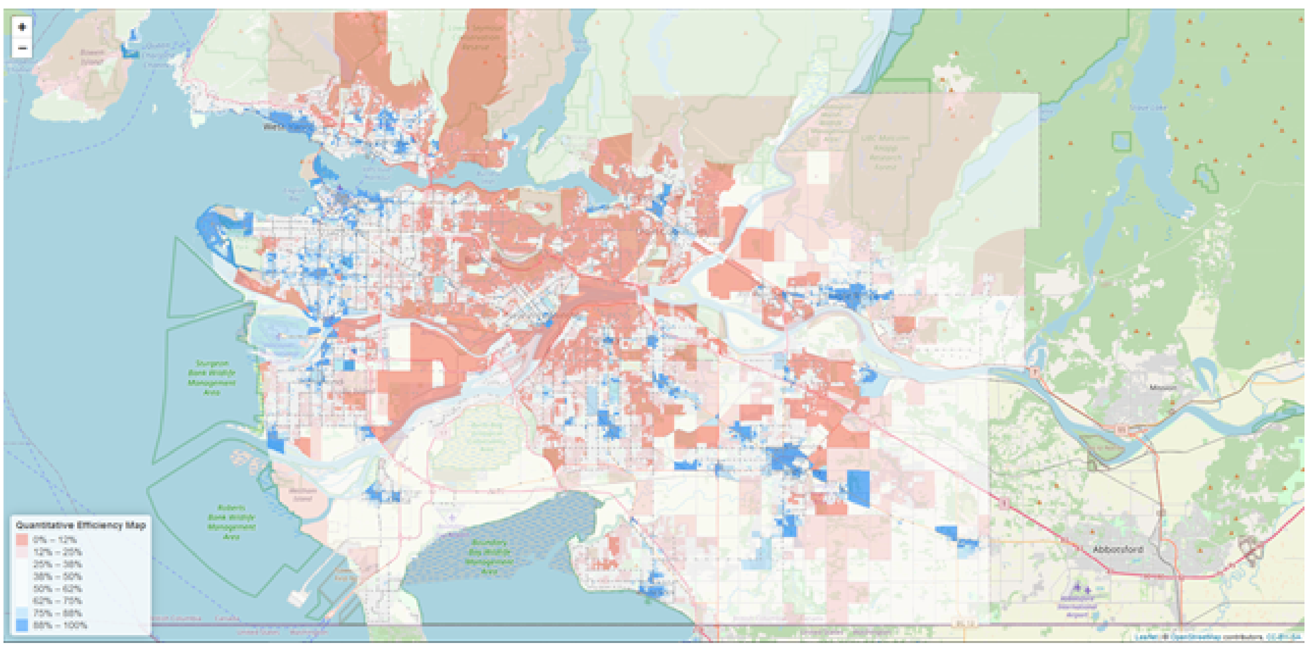


Figure 18. Efficiency choropleth using a discontinuous colour scheme.

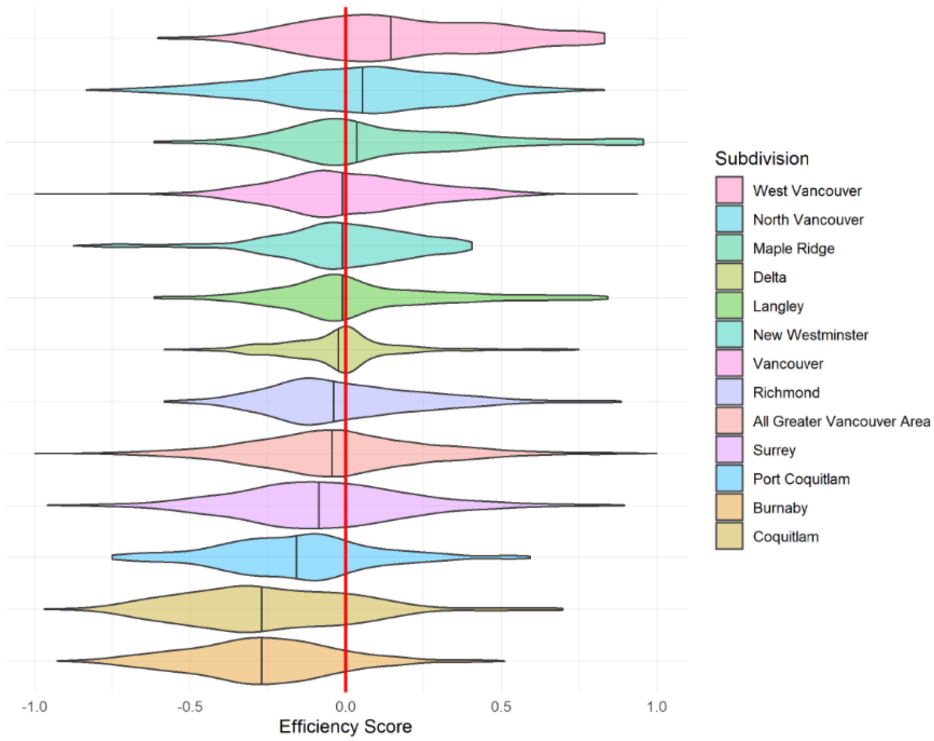


Figure 19. Violin plots of municipalities in the Greater Vancouver Area with the median efficiency score indicated for each municipality. An abline at an efficiency score of 0 separates municipalities that lack accessibilities and those with excess transit accessibility.